

**Thema der Dissertation:**

Wirksamkeit Individualpsychologisch-Psychoanalytischer Psychotherapie  
- eine Langzeitstudie zur Psychotherapieforschung.

**Dissertation**

zum Erwerb des Grades eines Dr. phil. im Fb 2 (Erziehungswissenschaft, Psychologie, Sport- und Bewegungswissenschaften)  
UNIVERSITÄT – GESAMTHOCHSCHULE – ESSEN-DUISBURG

**vorgelegt von:**

Diplom-Psychologe Dieter Schleussner  
Geburtsort: Gummersbach  
verantwortlich für die Kapitel 2, 10.3.5 - 10.4.5, 10.5 und 11 - 14

und

Diplom-Psychologe Franz Josef Klar  
Geburtsort: Köln  
verantwortlich für die Kapitel 1, 3 - 10.3.4 und 10.4.6 - 10.4.11

Datum der Disputation: 12.06.2003

Gutachter: Prof. Dr. Langenmayr  
Prof. Dr. Heisterkamp

**Die Wirksamkeit Individualpsychologisch-  
Psychoanalytischer Psychotherapie  
- eine Langzeitstudie zur Psychotherapieforschung**

## Inhalt

1.	Studienanlass: Der „Grawe-Schock“	1
2.	Überblick über Geschichte und Stand der empirischen Psychotherapieforschung, unter besonderer Berücksichtigung der Psychoanalyse	9
2.1	Phase 1, 1930 bis 1970: Ergebnisforschung	9
2.2	Phase 2, 1960 bis 1980: Zusammenhang zwischen Ergebnis und Prozess	15
2.3	Kritik an Effektstudien und experimentellem Gruppenvergleich	26
2.4	Phase 3, ab 1980: Leitthema „Feinstruktur des Therapieprozesses“ und tatsächliche Forschungspraxis	27
3.	Wissenschaftstheorie und Psychoanalyse	46
4.	Empiristisch-wissenschaftslogische Kriterien	51
5.	Vorentscheidungen	63
5.1	Begrenzung auf Individualpsychologisch-Psychoanalytische Psychotherapie	63
5.2	Trennung von Therapeut und Untersucher	63
5.3	Ausbildungsstand der Therapeuten	64
5.4	Krankheitsnachweis	64
5.5	Geographische Ausweitung der Untersuchung	64
5.6	Ethische Einwände hinsichtlich der Bildung der Kontrollgruppe	65
5.7	Langzeitstudie und Katamnese	65
5.8	Freiheit von den Finanzierungspartnern	66
5.9	Anonymität der Patienten	66
5.10	Handhabung der Rohwerte als objektive Daten	67
5.11	Stichprobenauswahl	68
5.12	Entscheidungsgründe für die angewandten Messverfahren	68
5.12.1	Fragen der Validität und Reliabilität	68
5.12.2	Der Gießener Beschwerdebogen / GBB	72
5.12.3	Der Minnesota Multiphasic Personality Inventory (Kurzform) / MMPI(K)	73
5.12.4	Der Veränderungsfragebogen des Erlebens und Verhaltens / VEV	74
5.13	Auswertungsvorentscheidungen	75
5.13.1	Festlegung der auszuwertenden Messzeitpunkte	75
5.13.2	Gründe für die Auswertung mittels t-Test und Varianzanalyse	75
6.	Studiendesign	77
7.	Finanzierung	78
8.	Studiendurchführung	79
8.1	Organisation des Untersuchungsablaufs	79
8.2	Reaktion/Beteiligung der Therapeuten	80
8.3	Stichprobenauswahl	80
8.4	Rücklaufquote	81
8.5	Datenqualität	82
8.6	Doppelte Dateneingabe/Überprüfung stichprobenartig	82

9.	Grundsätze bei der Hypothesenformulierung	82
10.	Ergebnisse der Effektivitätsstudie (Verlaufsstudie)	85
10.1	Die Anzahl der Teilnehmer zu den fünf Testzeitpunkten	85
10.2	Überblick über Mittelwerte und Standardabweichungen in den GBB-Skalen, den MMPI(K)-Skalen und im VEV der Untersuchungsgruppe (UG) und der Kontrollgruppe (KG) zu den verschiedenen Messzeitpunkten	87
10.3	Sechs Ergebnismaße für die Therapieeffektivität	90
10.3.1	Ergebnisse im VEV zu den Testzeitpunkten t2, t3, t4 und t5	91
10.3.2	Einschätzung der Veränderungen in der Untersuchungsgruppe durch Vergleich mit der Kontrollgruppe unbehandelter Patienten im und nach dem ersten Behandlungsjahr (Querschnittuntersuchung)	92
10.3.2.1	Prüfung der Strukturgleichheit von Untersuchungsgruppe (UG) und Kontrollgruppe (KG) als notwendige Voraussetzung des Gruppenvergleichs	93
10.3.2.2	t-Tests mit den Mittelwertunterschieden von Untersuchungsgruppe und Kontrollgruppe zu den Messzeitpunkten t1, t2 und t3 (Querschnittuntersuchungen)	101
10.3.3	Effektivitätsfeststellung durch Unterschiedsmessungen diverser Prä-Post-Vergleiche (Längsschnittuntersuchungen) mit Hilfe von GBB-Skalen, MMPI(K)-Skalen sowie durch VEV-Vergleiche im dreijährigen Untersuchungszeitraum	106
10.3.3.1	Prüfung der diagnostischen Strukturgleichheit von Untersuchungsgruppe (UG) und Kontrollgruppe (KG) als Voraussetzung der Aussagegültigkeit von Prä-Post-Vergleichen (Längsschnittvergleichen)	107
10.3.3.2	t-Tests zur Signifikanzprüfung der Mittelwertunterschiede der fünf Messzeitpunkte in der Untersuchungsgruppe (diverse Prä-Post-Längsschnittvergleiche)	112
10.3.3.3	t-Tests zur Signifikanzprüfung der Mittelwertunterschiede der drei Messzeitpunkte in der Kontrollgruppe (diverse Prä-Post-Längsschnittvergleiche)	118
10.3.4	Varianzanalyse zur Signifikanzprüfung des unterschiedlichen Verlaufs von UG und KG zwischen den Messzeitpunkten t1 und t3	122
10.3.5	Indirekte Veränderungsmessung: unsere Ergebnisse und die Ergebnisse von Vergleichsgruppen	128
10.3.5.1	Die Ergebnisse im MMPI(K)	128
10.3.5.2	Die Ergebnisse im GBB	134
10.3.5.3	Die Ergebnisse im VEV	144
10.3.6	Globalmaß der Therapieeffektivität (Effektstärke)	145
10.4	Varianzanalysen	153
10.4.1	Diagnosegruppen	154
10.4.2	Patientengeschlecht	156
10.4.3	Patientenalter	157
10.4.4	Dyaden Therapeut/Patient	159
10.4.5	Sitzungsfrequenz	161
10.4.6	Therapeutengeschlecht	164
10.4.7	Therapeutenberuf	165
10.4.8	Schulbildung des Patienten	167
10.4.9	Höhe des Anfangs-Lie	168
10.4.10	Anfänglicher Beschwerdedruck	170
10.4.11	Anfangsdepression	173
10.5	Zusammenfassung und Kommentierung der Effektivitätsstudie	176

11.	Die Katamneseuntersuchung	185
11.1	Einleitung	185
11.1.1	Zum Problem der Katamnese in der Psychoanalyse	185
11.1.2	Das Design unserer Katamnese studie	185
11.2	Auswertung der Katamnese daten	186
11.2.1	Wer sind die Katamnese teilnehmer?	187
11.2.1.1	Patienten der Untersuchungsgruppe und der Kontrollgruppe	187
11.2.1.2	Der Zeitpunkt ihres Therapieendes	188
11.2.1.3	Sozio-demografische Merkmale der Katamnese teilnehmer	189
11.2.2	Prüfung auf Unterschiede zwischen Teilnehmern und Nicht-Teilnehmern an der Katamnese	190
11.2.2.1	Altersverteilung	190
11.2.2.2	Geschlechterverteilung	191
11.2.2.3	Globalmaß der Therapieeffektivität (Effektstärke $t1:t5$ )	192
11.2.3	Zwei verschiedene Effektivitätsmaße für die Katamnese	192
11.2.3.1	Der Helping Alliance Questionnaire (HAQ)	192
11.2.3.2	Die Intrex–Kurzform	202
11.2.4	Drei Fragen zur Freude in der Therapie	214
11.2.5	Korrelationen	215
11.2.6	Zusammenfassung und Kommentierung der Katamnese ergebnisse	224
12.	Abschließende Stellungnahme zur Gesamtuntersuchung und Stellenwert der Ergebnisse im Kontext der psychoanalytischen Wirkungsforschung	228
13.	Aspekte zukünftiger Psychotherapieforschung	231
14.	Literaturverzeichnis	238

## ANHANG

Tabellen

Antworten auf die Freude-Frage

Verzeichnis der versandten Schreiben und Fragebögen

## 1. Studienanlass: Der „Grawe-Schock“

Den Anlass zu dieser Studie gab das „Forschungsgutachten zu Fragen eines Psychotherapeutengesetzes“ im Auftrag des Bundesministeriums für Jugend, Familie, Frauen und Gesundheit aus dem Jahre 1991, erstellt von Meyer, Richter, Grawe, Schulenberg und Schulte, üblicherweise als das „Grawe-Gutachten“ oder „Forschungsgutachten“ bezeichnet.

Die von Grawes Gutachten ausgehende Wirkung wäre wahrscheinlich eher gering und auf keinen Fall schockierend gewesen, wäre sein Gutachten nicht auf Nachfrage und durch Beauftragung aus dem politischen Bereich zustande gekommen. So aber verband sich für die Psychoanalyse mit dem Gutachten die Sorge hinsichtlich der Auswirkungen auf das damals in Vorbereitung befindliche Psychotherapeutengesetz. Diese Sorge war um so mehr begründet, weil das Gutachten mit seinen Ergebnissen an den Gesetzgeber die Empfehlung verband, bestimmte im Gutachten genannte Psychotherapieformen mangels wissenschaftlich nachgewiesener Wirksamkeit von der Versorgung der Bevölkerung im Sinne der Kassenzulassung auszuschließen.

Das Grawe-Gutachten verstand sich als eine Metaanalyse aller Therapieformen und bediente sich in der Bewertung unhinterfragt des Maßstabes des seit Jahrzehnten an allen deutschen Hochschulen im Bereich der Sozialwissenschaften und der Psychologie dominierenden empiristisch-positivistischen Wissenschaftsverständnisses und unterwarf von diesem Wissenschaftsverständnis ausgehend alle psychotherapeutischen Schulen wie selbstverständlich einer empirisch-positivistischen Effektivitätsbewertung. Von daher war es denn auch nicht sehr überraschend, dass die Verhaltenstherapie, die ihr Wissenschaftsverständnis aus dem Positivismus ableitet, in der Begutachtung Grawes deutlich bessere ‚Noten‘ erhielt als die analytischen Schulen. Der Grundgedanke des Positivismus, nämlich der des ‚Zählens und Messens‘ von im empirischen Sinne positiv beobachtbaren Tatsachen, steht dem hermeneutisch beschreibend-verstehenden Wissenschaftsverständnis der analytischen Schulen so entgegen, dass diese Schulen sich von ihrem Denkansatz geleitet nie „um die in ihren Augen zumeist mit fragwürdiger Methodik durchgeführten Erfolgsuntersuchungen der akademischen Psychologen“ (Mertens, 1994, S. 40) gekümmert hatten, - im übrigen ja auch scheinbar nicht kümmern mussten, weil sie in den 60er Jahren ja als das erste psychotherapeutische Wirkverfahren Kassenzulassung erhalten hatte.

Was die psychoanalytische Psychotherapie anbelangt, so kommt Grawe trotz des von ihm konstatierten weitgehenden Fehlens empirischer Wirksamkeitsnachweise letztlich dann doch sehr überraschend zu einem vergleichsweise milden Urteil, wenn er der psychoanalytischen Therapie eben nicht Wirkungslosigkeit bescheinigt, sondern seine Kritik lediglich gegen die analytischen Psychotherapeuten richtet, denen er zum einen Unfähigkeit, ihr Instrumentarium wirkungsvoll einzusetzen und zum anderen Ignoranz, die nötigen Wirksamkeitsnachweise beizubringen, vorwirft. Hierzu ein Zitat von Grawe selber:

„Der wissenschaftliche Entwicklungsstand zur psychoanalytischen Therapie ist allerdings insgesamt positiver einzuschätzen, als es aus diesem Ergebnisbericht hervorgeht, der sich auf Studien beschränkt, die stichhaltige Aussagen über die Wirksamkeit psychoanalytischer Therapie zulassen. Neben den hier referierten kontrollierten Wirkungsuntersuchungen gibt es eine stetige Produktion an Studien zu Wirkfaktoren ... mit einem vor allem in den letzten Jahren angestiegenen methodischen Niveau, die auch den strengen Prüfkriterien angesehener nicht-psychoanalytischer Fachzeitschriften standhalten. Die wissenschaftliche Auseinandersetzung mit der psychoanalytischen Therapie findet also inzwischen nicht mehr nur innerhalb psychoanalytischer Zirkel statt nach Kriterien, die von einer Mehrzahl von Wissenschaftlern nicht akzeptiert werden, sondern erfolgt inzwischen wenigstens zum Teil nach den allgemeinen Standards empirisch-wissenschaftlicher Arbeit und ist dadurch in den allgemeinen wissenschaftlichen Austauschprozess einbezogen.“ (Grawe, 1994, S. 197)

Und noch einmal Grawe:

„Die übliche psychoanalytische Praxis ist daher sehr viel weniger wissenschaftlich fundiert als der psychoanalytische Ansatz als solcher. Die dazu vorliegenden Ergebnisse legen eine erhebliche Revision der psychoanalytischen Veränderungstheorie nahe, die in vielen Punkten als widerlegt angesehen werden kann. Solche Veränderungen sind aber der normale Lauf der Wissenschaft und kein Grund, dem psychoanalytischen Therapieansatz die Wissenschaftlichkeit abzusprechen.“ (Grawe, 1994, S. 741)

Trotz dieser für den psychoanalytischen Ansatz letztlich noch glimpflichen Einschätzung nahm der damalige Vorsitzende der DGPT Prof. Dr. U. Streeck die Situation zum Anlass, an die der DGPT angeschlossenen Institute ein besonders dringliches und besorgtes Rundschreiben zu verfassen, in dem er warnend auf die Konsequenzen hinwies, dass die Grawe'schen Arbeiten schon bald von Kostenträgern, kassenärztlichen Vereinigungen u.a. gegen die Vertreter der Psychoanalyse verwendet werden könnten. Er bat „nachdrücklich darum, sich dieser Thematik anzunehmen“ und „die Ergebnisse dem geschäftsführenden Vorstand für dessen Arbeit und für die zu erwartenden berufspolitischen Auseinandersetzungen zur Verfügung zu stellen“. (Quelle: Brief von 1992, vgl. Anhang)

Besonders hart traf das „Grawe-Gutachten“ die im Forschungsgutachten (unabsichtlich/absichtlich, auf jeden Fall aber missverständlich) als „Individualtherapie“ bezeichnete „Individualpsychologisch-Psychoanalytische Psychotherapie“, der gegenüber Grawe nicht die gleiche Milde wie der Psychoanalyse der freudianischen Gesellschaften entgegenbrachte, - offensichtlich u.a. auch deshalb nicht, weil er die „Individualpsychologisch-Psychoanalytische Psychotherapie“ nicht als psychoanalytisches Verfahren wahrnahm/wahrnimmt bzw.

wahrnehmen wollte/will. Zu den Wirksamkeitsnachweisen der individualpsychologischen Psychotherapie führt Grawe aus:

„Die beiden Studien, in denen die Wirksamkeit von Individualtherapie geprüft wurde ... weisen erhebliche Mängel auf. Man erfährt nur, dass es sich um ambulante Einzeltherapie gehandelt hat. Die Art der behandelten Patienten wird nicht näher spezifiziert. Außerdem wurde nur ein einziges Therapieerfolgsmaß erhoben, nämlich die Annäherung von Selbst- und Idealbild. Eine solche Annäherung konnte im Prä-Post-Vergleich in beiden Studien statistisch gesichert werden. Es wurde aber kein statistischer Vergleich mit den Veränderungen in der Kontrollgruppe vorgenommen. Gegenüber der mit der Individualtherapie verglichenen Gesprächspsychotherapie fanden sich keine signifikanten Unterschiede in dem einzigen Erfolgsmaß. Die Ergebnislage zur Individualtherapie ist daher quantitativ und qualitativ außerordentlich dürftig und liegt weit unterhalb der Grenze, ab der die Wirksamkeit eines Therapieverfahrens als ausreichend gesichert gelten kann. Zum gegenwärtigen Zeitpunkt kann daher die Individualtherapie für sich nicht in Anspruch nehmen, zu den empirisch bewährten Therapieverfahren gezählt zu werden.“ (Forschungsgutachten, S. 86)

Was die Anzahl der vorliegenden empirischen Wirksamkeitsuntersuchungen zur individualpsychologischen Psychotherapie anbelangt, so muss man Grawe zumindest zugestehen, dass es in diesem Punkt tatsächlich recht dürftig aussieht. Bei den beiden von Grawe erwähnten Studien handelt sich um die Untersuchungen mit der von ihm vergebenen Studiennummer 64006 und 68051. Im Anhang des Psychotherapiegutachtens findet man unter 64006: „Shlien, J.M., Comparison of results with different Forms of Psychotherapy, Amer. J. Psychotherapy, 1964, 18.15-22. Und unter 68051 findet man: Butler, J.M., Self-Ideal Congruence in Psychotherapy, Psychother: Theory, Res.Practice, 1968,5,13-17. Hierbei handelt es sich offensichtlich um zwei aus dem amerikanischen Raum kommende Untersuchungen, deren Übertragbarkeit auf den deutschsprachigen Raum unabhängig von Grawes Kritik sowieso bezweifelt werden kann. So gesehen können diese beiden Studien zwar die „Individualpsychologisch-Psychoanalytische Psychotherapie“ nicht belasten, doch reduziert sich die für den deutschsprachigen Raum vorliegende Untersuchungsmenge damit auf null, was die Lage für die „Individualpsychologisch-Psychoanalytische Psychotherapie“ nicht besser macht.

Wohl infolge der schockartigen Wirkung hat es dann kurze Zeit nach der Veröffentlichung des Forschungsgutachtens und anderer Publikationen von Grawe eine ganze Reihe von kritischen Er widerungen gegeben, so z.B. von Kaiser (1993), Leikert (1994), Brockmann (1995), Kächele (1995). Eine der wohl umfangreichsten Er widerungen ist die Publikation von Mertens (1994) „Psychoanalyse auf dem Prüfstand? Eine Er widerung auf die Meta-Analyse von Klaus Grawe“. Mertens wirft Grawe einen „Imperialismus einer einzigen Methodologie und Wissenschaftsauffassung“ (S. 82) mit „fundamentalistischem Anspruch“ (S. 83) vor, betont dem gegenüber die Besonderheit des psychoanalytischen Gegenstandes und leitet hieraus die Schlussfolgerung ab, dass die ‚Methodik der Empirischen Sozialforschung‘ ungeeignet sei, den Gegenstand psychoanalytischer Therapie überhaupt geeignet erfassen zu können. Er fordert im Gegenzug, dass „jede Wissenschaft



mit den ihr adäquaten Methoden ihrer eigenen wissenschaftlichen Matrix arbeiten“ können müsse, „ohne sich damit dem Verdacht der Unwissenschaftlichkeit auszusetzen“ (Mertens, 1994, S. 13). Mertens betont das Zubilligungsverhältnis der ‚Angemessenheit von Untersuchungsgegenstand und Untersuchungsmethodik‘ als ein wissenschaftstheoretisches Grundkriterium aller Wissenschaftstheorien und leitet von da aus ab, dass es keine „einheitswissenschaftliche Methodologie“ (S.13) geben könne, weshalb die psychoanalytischen Vertreter Vorsicht walten lassen sollten, aus berufspolitischer Eilfertigkeit heraus der Versuchung zu erliegen, „etwas vorschnell mit fragwürdigen Methoden und Kriterien messen zu wollen“ (Mertens, 1994, S.16 f.). Mertens verwahrt sich gegen den Vorwurf, die Psychoanalyse sei deshalb keine Wissenschaft, weil sie keine „Empirie“ und kein „Falsifikationstheorem“ kenne (Mertens, 1994, S. 83) und sich dadurch gegen eine kritische und intersubjektive Überprüfung immunisiere. Beide Wissenschaftskriterien seien in der Psychoanalyse sehr wohl vorhanden, jedoch anders hergeleitet. Empirie sei in der psychoanalytischen Gegenstandsbildung eine beim Therapeuten ausgelöste und von diesem real wahrnehmbare (also empirisch objektiv vorhandene) Gefühlsreaktion als Antwort auf das beobachtbare Verhalten des Patienten. Diese Gefühlsreaktion des Therapeuten sei nach den Regeln der tiefenhermeneutischen Methode beschreibbar, am Fall überprüfbar und damit also auch „falsifizierbar“. Leider wird die hier in Ansätzen vorfindbare Gegenstandsbildung einer eigenständigen psychoanalytischen Wissenschaftstheorie von Mertens nicht weiter vorangetrieben. Statt dessen setzt er sich mit den nur randläufig wissenschaftstheoretische Problembereiche streifenden Aspekten der Aussagen von Grawe kritisch auseinander, und zwar, ob die Verhaltenstherapie oder die Psychoanalyse die schwerer gestörten Patienten behandle, wie die von den 50er bis in den 80er Jahren durchgeführte „Menninger-Studie“ zu interpretieren ist, über das Vorhandensein oder Nichtvorhandensein und die theoretische Würdigung „supportiver“ Therapiemaßnahmen in der psychoanalytischen Therapie, mit der Frage nach dem therapeutischen Nutzen von psychoanalytischen Langzeitpsychotherapien angesichts „rascher Anfangserfolge“ von Verhaltenstherapien, mit den von Grawe behaupteten angeblich schädlichen Auswirkungen der abstinenter Lehranalyse auf die Schulung der therapeutisch relevanten zwischenmenschlichen Interaktionsfähigkeit des späteren Therapeuten sowie mit den ebenfalls behaupteten angeblichen negativen Auswirkungen der außeruniversitären und damit der öffentlichen Korrektivwirkung nicht zugänglichen analytischen Ausbildung auf die wissenschaftliche Erkenntnisbereitschaft der Psychoanalytiker. Schließlich wird auch noch der Versuch Grawes, die analytische Therapie als eine Teilmenge der behavioristischen Therapie zu sehen, durch Umkehrung konterkariert. Weiter wird die Forderung Grawes nach einer allgemeinen Einheitspsychotherapie kritisch hinterfragt, Grawes Vorwurf einer angeblich von den analytischen Therapieschulen blockierten wertneutralen Indikationsstellung und wirkungsorientierten Überweisung von Patienten an nichtanalytische Therapeuten diskutiert und die Verhinderung der Möglichkeit eines friedlichen Nebeneinanders von Verhaltenstherapie und Psychoanalyse angesichts immer erneuter „Postulierung einer Überlegenheit der kognitiv-behavioralen Therapie über die Psychoanalyse seitens Grawe et al“ (Mertens, 1994, S. 79) bedauert. Alle diese von Mertens stellvertretend für die gesamte analytische Therapiebewegung vorgetragenen Argumentationspunkte haben letztlich das Ziel, die von Grawe geforderte Notwendigkeit eines empiristisch verstandenen Wirksamkeitsnachweises der analytischen Psychotherapien zu verneinen und die Beibringung eines solchen Nachweises unnötig erscheinen zu lassen. Und er warnt, „... psychometrische Verfahren und Evaluationstechniken

zur Bestimmung des Psychotherapieerfolgs nicht einfach unbesehen zu übernehmen“ (Mertens, 1995, S. 405).

Trotz der theoretisch eher ablehnenden Erwiderungen häufen sich nach Grawe die Anzahl der in Fachzeitschriften veröffentlichten neueren Untersuchungen zur Wirksamkeit von psychoanalytischen Psychotherapien. In diesem Zusammenhang sei auf Groß und Riedel (1995), Heuft u.a. (1996), Heinzel u.a. (1996), Leuzinger-Bohleber u.a. (2001), Sandell u.a. (2001) sowie auf den Sammelband „Langzeitpsychotherapie“ von Stuhr, Leuzinger-Bohleber und Beutel (2001) verwiesen, in dem gleich neun Studien verschiedener psychoanalytischer Forscher Darstellung finden. Es ist auffallend, dass diese Untersuchungen durchweg eine Vermischung zwischen analytischer Einzelfallbeschreibung und Einbeziehung von methodischen Aspekten des empirisch-positivistischen Verfahrensinstrumentariums liefern. Bei aller Betonung des Integrationswillens ist bei der Lektüre dieser Studien immer eine gewisse Halbherzigkeit durchzuspüren. Man hat den Eindruck, als solle Wissenschaftlichkeit demonstriert werden, um Legitimationskritik aus dem „empirischen Lager“ vorbeugend zu begegnen. Der Schock scheint offensichtlich trotz aller theoretisch teilweise vernichtender Auseinandersetzung mit Grawe bei den Vertretern der analytischen Therapie ein permanent wirkendes Unbehagen hinterlassen zu haben, das nachhaltig abzubauen vielleicht nur dadurch möglich zu sein scheint, dass in irgend einer Weise doch noch der geforderte „empirische“ Wirksamkeitsnachweis nachgeliefert wird. Denn es ist klar, würde der geforderte empirische Wirksamkeitsnachweis beigebracht, dann wäre die wissenschaftstheoretische Auseinandersetzung über die Notwendigkeit solcher Nachweise zumindest insofern überholt, als dass der geforderte Nachweis dann ja im positivistischen Sinne - völlig gleichgültig ob nun mit einem positiven oder negativen Ergebnis - gleichwohl „vorliegend“, also „positiv“ beigebracht wäre. Vielleicht könnten dann, sozusagen im Angesicht eines dann vorliegenden Wirksamkeitsnachweises, die wissenschaftstheoretischen Argumente über das Für und Wider eines solchen Nachweises überhaupt erst respektvolles Gehör finden.

Parallel zu den sich häufenden Wirksamkeitsforschungen gibt auch eine ganze Reihe von Autoren, die zu den empirischen Bemühungen die theoretische Begleitung liefern und für ein „sowohl als auch“ des methodischen Vorgehens plädieren. Zu nennen ist hier u.a. Stuhr mit seinem bereits Programm umschreibenden Aufsatz „Methodische Überlegungen zur Kombination qualitativer und quantitativer Methoden in der psychoanalytischen Katamneseforschung und Hinweise zu ihrer Integration“ (Stuhr, 2001) oder auch Hermann Faller (1994), der schreibt:

„Wie in gegenstandsangemessener Forschung qualitatives und quantitatives Vorgehen einander nicht ausschließen müssen, so gilt das auch für psychoanalytische Therapie und empirische Forschung. Mehr noch, psychoanalytische Therapieforschung sollte die von der psychologischen Forschung zur Verfügung gestellten Designs und Methoden nutzen und kreativ weiterentwickeln (vgl. Leuzinger-Bohleber u. Kächele 1990). Um beiden Aspekten, dem Qualitativen und dem Empirischen, zu ihrem Recht zu verhelfen, ist Kreativität gefragt. Dabei sind

Berührungspunkte in beiden Richtungen wenig hilfreich. Beide Forschungstraditionen zu integrieren, erfordert eine kreative Leistung, die Neues zu schaffen in der Lage ist, um bisherige Beschränkungen zu überwinden. Deshalb ist qualitative Forschung zu weiten Teilen noch Desiderat, und es ist vielleicht angebrachter, von einem Forschungsprogramm zu sprechen, das der Realisierung harret.“ (Faller, 1994)

Dagegen hält Kaiser solche Kompromisse zwischen der quantitativen Psychotherapieforschung und der psychoanalytischen Forschung für eine „(pseudo)-wissenschaftliche Aktivität“ (Kaiser, 1993, S. 359):

„Die Bezeichnung ‚psychoanalytisch orientierter Psychotherapieforscher‘ oder Überschriften wie ‚Psychoanalytische Psychotherapieforschung‘ (Kächele 1992) suggerieren, es gäbe zwischen quantitativer Psychotherapieforschung und psychoanalytischer Forschung ein Drittes. Ich halte das für Etikettenschwindel.“ (Kaiser, 1993, S. 350)

Kaisers Argumentation ist insofern nicht ganz richtig, als er unterstellt, dass es die „quantitative Psychotherapieforschung“ bereits hinreichend umfangreich gibt.

Tatsache ist, dass es zwar den „ersten Weg“ der psychoanalytischen Forschung, - und im Anschluss an Grawe nun in zunehmend umfänglicher Ausweitung auch den von Kaiser bezeichneten „Dritten Weg“, nämlich den Versuch einer quantitativen Beimischung zu psychoanalytischen Forschungsansätzen gibt, doch den implizit mitgedachten sogenannten „zweiten Weg“, nämlich den der Anwendung quantitativer Methodik auf psychoanalytische Psychotherapien, den gibt es bisher eigentlich nicht. Denn wenn Grawe ein Verdienst hat, dann ist es sicherlich unbestritten dieses, dass er ein umfangreiches Literaturstudium absolviert und festgestellt hat, dass es den „Zweiten Weg“, nämlich die im empirischen Wissenschaftsverständnis methodologisch saubere psychotherapeutische Wirkungsforschung - bis auf die von ihm genannten und methodisch kritisierten Ausnahmen - so gut wie nicht gibt. Und genau in diese Lücke zielt die nachfolgende und von uns vorgelegte Effektivitätsstudie. Von Anfang an galt nämlich für uns die methodologische Grundentscheidung, keinen „Dritten Weg“ gehen zu wollen, sondern ganz im Gegenteil, den empirisch-nomothetischen Ansatz möglichst in „Reinkultur“ auf die psychoanalytische Psychotherapie anwenden zu wollen und damit eine auf empiristischer Grundlage basierende Aussage über Wirksamkeit oder Unwirksamkeit dieser Therapieform vorlegen zu können.

An diesem Punkt der Überlegungen stellte sich uns hinsichtlich der „Individualpsychologisch-Psychoanalytischen Psychotherapie“ zunächst die konkrete Frage, ob, und wenn ja, welche Untersuchungen zur individualpsychologischen Psychotherapie denn seit Grawe vorgelegt wurden. Zu nennen sind hier die im „geschichtlichen Teil“ dieser Arbeit noch ausführlicher zu referierenden Arbeiten von Rogner (1994, vgl. unten), die Veröffentlichung von Heinzel u.a. (1996, vgl. unten) als auch die Untersuchung von Brockmann u.a. (2001, vgl. unten), - die beiden letztgenannten nur mit Einschränkungen (vgl. unten) :

Rogners Untersuchung begrenzt sich ausschließlich auf „Individualpsychologisch-Psychoanalytische Psychotherapien“. Er vergleicht zunächst eine Patientengruppe nach einjähriger Psychotherapie mit einer Kontrollgruppe (bestehend aus am Anfang ihrer Therapie stehende Patienten) und bedient sich hierbei verschiedener psychometrischer Testverfahren. In einem zweiten (katamnestischen) Teil sollten die Patienten ein Jahr nach Abschluss der Therapie den Umfang der Veränderung einschätzen. In beiden Untersuchungsabschnitten konnten deutliche Therapieerfolge gemessen werden. Allerdings kann der Untersuchung - gemessen an den Standards der Methodologie der empirischen Sozialforschung - das Fehlen einer ‚quasi experimentellen Versuchsanordnung‘ vorgehalten werden. Rogner hat versucht, die Schwierigkeiten der Bildung einer Kontrollgruppe durch eine Hilfskonstruktion zu umgehen, indem er eine am Anfang ihrer Therapie stehende Patientengruppe wie eine Kontrollgruppe handhabte und diese dann mit einer katamnestisch befragten Patientengruppe verglich.

Heinzel u.a. legten 1996 eine eigene Untersuchung zur Wirksamkeit von Psychotherapie vor, deren Datenmaterial sich u.a. auch auf Therapeuten der DGIP (Deutsche Gesellschaft für Individualpsychologie) stützt. Von 1979 Mitgliedern der beiden Verbände DGPT und DGIP brachte die DGIP lediglich 275 Mitglieder ein (S. 22), wobei der genaue Anteil der in die Untersuchung tatsächlich eingeflossenen individualpsychologischen Psychotherapien noch weit geringer zu sein scheint, weil aus der Grundgesamtheit lediglich eine Zufallsstichprobe von nur 394 Therapeuten/-innen zur Mitarbeit aufgefordert wurden, wobei die genaue Anzahl der darin enthaltenen DGIP-Mitglieder nicht genannt wird. Bei dieser Untersuchung handelt es sich um eine rein katamnestische Studie, die bereits deshalb schon vom methodischen Ansatz her gesehen nicht allen Standards eines „empirischen Wissenschaftsverständnisses“ genügt.

Über die von Brockmann u.a. durchgeführte Untersuchung gibt es derzeit leider nur eine eher knapp zu nennende Veröffentlichung (Brockmann, 2001, S. 271-276). Es handelt sich um eine Untersuchung, in die neben individualpsychologischen Psychotherapien vergleichend auch die Ergebnisse verhaltenstherapeutischer Langzeitpsychotherapien mit einbezogen wurden, was die Autoren als integrativen Fortschritt preisen. Die Untersuchung möchte eine schulübergreifend-integrative Grundhaltung vertreten. Gemeint ist, dass die Untersuchung der Psychotherapiewirksamkeit nicht mehr in Bezug zum angewandten Therapieverfahren, sondern in Bezug zur Persönlichkeit des Therapeuten gesehen wird. Gegen diesen Ansatz ist vom Prinzip nichts einzuwenden, - nur kann die „Individualpsychologisch-Psychoanalytische Psychotherapie“ aus den Ergebnissen einer so angelegten Studie eben nicht mehr den Nachweis ihrer schulrichtungseigenen Wirksamkeit ableiten, den sie in der berufspolitischen Auseinandersetzung doch so dringend bedarf. Gleichwohl scheint sich diese Untersuchung um ein im Sinne der „Methodik der empirischen Sozialforschung“ ‚sauberes‘ Design zu bemühen: Befragung mit objektivierbaren Fragebögen (SCL-90-R, IIP, GAS, GT und VEV) bei Behandlungsbeginn, mehreren erneuten Befragungen nach bestimmten Zeitabständen und eine katamnestische Wiederbefragung dreieinhalb Jahre nach Therapie-Ende sowie Auswertung der Daten mit Hilfe von Varianzanalysen, mit deren Hilfe man signifikante Therapieeffekte gemessen haben will. Nähere und überprüfbare Angaben hierzu findet man in dieser Veröffentlichung jedoch leider nicht.

Warum gibt es so wenige Untersuchungen aus dem Bereich der „Individualpsychologisch-Psychoanalytische Psychotherapie“? Heinzl u.a. (1996) haben hierzu eine folgende Meinung:

„Warum spezifische Wirksamkeitsnachweise bisher so schwer zu erbringen sind, liegt u.E. im wesentlichen nicht so sehr an der mangelnden Differenziertheit oder anderen Unzulänglichkeiten des Designs, oder an zu wenig eindeutigen Fragestellungen oder Kategorien. Wir glauben, dass die Ursache im fundamentalen Verständnis von Wissenschaft liegt, welches in unseren Breiten noch immer hauptsächlich an den Naturwissenschaften orientiert und damit vorwiegend positivistisch ist, oder – mit Klaus Holzkamps (1970, 1989) Worten formuliert – einem ‚naiven Empirismus‘ unterliegt. Ein großer Teil der in den letzten Jahren veröffentlichten Untersuchungen ist offenbar noch von diesem positivistischen Paradigma und dem auf ihm basierenden Falsifikationstheorem von Popper getragen.“ (Heinzl u.a., 1996, S. 14)

Heinzl verweist hier darauf, dass die psychoanalytischen Schulen es bisher nicht vermocht haben, sich von der empiristisch-positivistischen Wissenschaftsdominanz zu befreien und stattdessen von dieser Dominanz geradezu gelähmt darauf verzichten, sich mit der Frage der Wirksamkeit ihrer Therapien und einer nicht-neurotischen Form zu beschäftigen. Man merkt den in den letzten Jahren vorgelegten Untersuchungen an, dass sie sich unter dem von Grawe ausgehenden Druck notgedrungen, aber nicht von Überzeugung getragen, zu Konzessionen an eine nicht wirklich angeeignete Methodologie bereitgefunden haben. So verständlich dies unter dem politisch erzeugten Legitimationsdruck auch sein mag, letztlich sind diese Ansätze zweiseitig inkonsequent, - inkonsequent dem psychoanalytischen aber auch inkonsequent dem empiristischen Ansatz gegenüber.

## 2. Überblick über Geschichte und Stand der empirischen Psychotherapieforschung, unter besonderer Berücksichtigung der Psychoanalyse

Die Geschichte der Psychotherapieforschung wird vielfach in Phasen eingeteilt, so auch bei Kächele (1992) sowie Fischer und Klein (1997). Die Phaseneinteilung geht zurück auf Shapiro (1990, zit. bei Fischer und Klein), der drei „Generationen“ oder Epochen der Psychotherapieforschung ausgemacht hat. Die erste sei bestimmt durch die *Ergebnisforschung*, zunächst um den Nutzen einer Therapie in Fachkreisen und der Öffentlichkeit unter Beweis zu stellen. Die zweite Epoche sei gekennzeichnet durch das Interesse am *Zusammenhang zwischen Ergebnis- und Prozessvariablen*. In der dritten Phase schließlich richte sich das Interesse auf die *Mikrodynamik*, die Feinstruktur des Therapieprozesses. Wie Epochen der Kulturgeschichte ließe sich diese Phasenfolge beobachten, und zwar innerhalb einer bestimmten Schulrichtung ebenso wie in der Psychotherapieforschung insgesamt.

Kächele (1992) unterteilt und beschreibt für die psychoanalytische Forschung die drei Phasen wie folgt:

1. Phase, beginnend in den dreißiger Jahren, dominierend in den Jahren 1950-1970. Leitfrage: „Nützt Psychotherapie bzw. Psychoanalyse überhaupt?“
2. Phase, dominierend von 1960-1980 mit der Leitfrage: „Was muss im Verlauf geschehen, dass am Ende ein positives Ergebnis erwartet werden kann?“
3. Phase, die in der zweiten Phase festgestellte Komplexität des Therapiegeschehens hat nun zur Konsequenz, die Untersuchung des Therapieprozesses zu intensivieren. Die Leitfrage lautet: „Welche mikroprozessualen Vorgänge konstituieren die klinisch konzipierbaren Makroprozesse?“

Wir folgen nun der Einteilung von Kächele und beginnen mit der 1. Phase:

### 2.1 Die Phase 1 der empirischen Psychotherapieforschung

Bachrach (1991) gibt einen Überblick über frühe statistische Untersuchungen:

- 1) **Coreat**, 1917, berichtete über Therapieergebnisse von 93 Fällen. Sie basierten auf seinen eigenen persönlichen Untersuchungen und Erfahrungen. Die Verbesserungen waren fast gleich für alle Diagnosekategorien, die schwierigen Fälle brauchten eine bis sechs Monate andauernde Behandlung.
- 2) **Fenichel** berichtete 1930 über Analysen, die unter der Obhut des Berliner Psychoanalytischen Institutes zwischen 1920 und 1930 durchgeführt wurden. Diagnostisch handelte es sich zum großen Teil um Neurosen, zum kleineren Teil um Psychosen. 60 % der Neurosen und 23 % der psychotischen Fälle wurden als substantiell verbessert eingeschätzt, 22 % der Neurosen und 24 % der Psychosen wurden als unverändert oder schlimmer beurteilt.
- 3) **Jones** berichtete 1936 über Ergebnisse aus der Londoner Psychoanalytischen Klinik von 1926-1936. Von 738 Bewerbern wurden 74 in Psychoanalyse genommen (80 % Neurotiker, 20 % Psychotiker). 47 % der

Neurosefälle wurde als substantiell verbessert beurteilt, 10 % als unverändert oder schlimmer, alle außer einem der 15 psychotischen Fälle waren Misserfolge.

4) **Alexander** berichtete 1937 über die Arbeit der Chicago Psychoanalytischen Klinik von 1932 bis 1937. Von den 157 Fällen, die in Alexanders Bericht enthalten waren, waren 22 % als neurotisch, 5 % als psychotisch klassifiziert. Die übrigen erhielten psychosomatische Diagnosen. 63 % der Neurotiker, 40% der Psychotiker und 77 % der Psychosomatiker profitierten substantiell von ihrer Therapie, weniger als 10 % von allen blieben unverändert oder verschlimmerten sich.

5) **Kessel und Hyman** berichteten 1933 von 21 psychoneurotischen und 8 psychotischen Fällen. 89% der neurotischen Fälle profitierten substantiell von der Behandlung (6 % veränderten sich nicht oder wurden schlimmer). Alle psychotischen Patienten veränderten sich nicht oder wurden schlimmer.

6) In einem Überblicksartikel berichtete **Knight** 1941 von 100 (31 psychoneurotische, 8 psychotische und 31 anders klassifizierte) Fällen, psychoanalytisch behandelt an der Menninger Klinik zwischen 1932 und 1942. 80 % der Neurotiker und 40 % der Psychotiker veränderten sich positiv, nur 1 % der neurotischen und 42 % der psychotischen Fälle blieben unverändert oder waren schlimmer geworden. Ein Überblick über alle bis dahin vorgelegten Untersuchungen zeigte eine Verbesserungsrate von 56-63 % für die nicht psychotischen Fälle, von 78 % für die psychosomatischen Fälle und von 25 % für die Psychosen. Knight fügte dann diesen Ergebnissen wichtige Überlegungen an, z.B. zu der Notwendigkeit, die Analysen gut ausgebildeter erfahrener Analytiker zu evaluieren, die Patienten behandeln, für die sie besonders von ihrer Veranlagung und spezifischen Erfahrung geeignet sind. In diesem Zusammenhang beklagte Fenichel 1945 schon das Fehlen von standardisierten operationalen Definitionen.

1942 folgte aus dem Berliner Institut von **F. Böhm** ein Bericht über 419 abgeschlossene psychoanalytische Behandlungen, den H. Duehrssen (1962) erwähnt.

Versucht man die Studienergebnisse zusammenzufassen, so könnte man sagen, dass die Wirksamkeit der Psychoanalyse festgestellt wurde, und zwar bei den neurotischen Fällen von 60% bis zu über 80% der behandelten Fälle, für die psychosomatischen Störungen waren die Erfolgsraten z.T. höher, für die psychotischen Fälle lagen sie erheblich niedriger.

Große Forschungsaktivität wurde in dieser Phase ausgelöst durch **Eysenck (1952)**, der behauptete, dass die Erfolgsrate von Psychotherapie nicht höher sei, als die spontane Remissionsrate. Eysenck stützte sich auf einen Vergleich der Besserungsrate von 24 Erfolgsberichten mit den spontanen Remissionsraten aus zwei Studien, die den natürlichen Verlauf psychischer Störungen untersuchten. Obwohl schon gleich nach Erscheinen des Artikels eine Fülle von stichhaltigen Einwänden gegen die Richtigkeit der Zahlen selbst und ihre Interpretation durch Eysenck vorgebracht wurden (z.B. Dührssen und Jorswieck, 1965), hatte die Studie einen nachhaltigen Einfluss auf die Entwicklung der Psychotherapieforschung.

Noch in der ersten Nummer des im Jahre 1991 erschienenen Journals „Psychotherapy Research“ gibt es einen Artikel, der sich ausschließlich mit der Interpretation der Zahlen aus Eysenck's 1952 erschienenen Artikel beschäftigt (Grawe, 1992). Darin wird überzeugend dargelegt, dass Eysenck's Zahlen schon damals

nichts weiter bedeutet haben, als dass es einem psychotherapeutisch behandeltem Patienten schon nach zwei Monaten so gut geht, wie einem unbehandeltem erst nach zwei Jahren (McNeilly und Howard, 1991).

Im deutschen Sprachraum wird diese erste Phase der Forschung durch die Studie von **Dührssen & Jorswieck (1965)** repräsentiert. Sie belegten die Leistungsfähigkeit von analytischer Psychotherapie mit einer mittleren Dauer von ca. 100 Stunden. In dieser Studie wurden die Krankenhausaufenthalte in der Poliklinik der AOK Berlin (jeweils im Jahrfünft vor und nach der Therapie) von Patienten, die psychoanalytisch behandelt wurden, untersucht. Es wurden die Daten über Krankenhausaufenthalte von drei Personengruppen verglichen:

- Patienten, die psychoanalytisch behandelt worden waren.
- Patienten, die nur zu einem Erstgespräch gekommen waren, aber wegen der Warteliste nicht behandelt wurden.
- Versicherte der AOK, die keinen Kontakt zur Poliklinik gehabt haben.

Es wurden nun die Krankenhaustage in den fünf Jahren vor und in den fünf Jahren nach der Therapie bzw. dem Erstgespräch und bei den übrigen Versicherten in dem entsprechenden Zeitraum vor und nach einem Stichtag gezählt. Dabei wurden alle Krankenhaustage mitgezählt, gleichgültig, ob die Einweisungsdiagnose eine organische oder eine neurotische Erkrankung vermuten ließ.



Die Ergebnisse waren deutlich (siehe Tabelle 1).

**Tabelle 1: Kumulierte Krankenhausaufenthalte (aus Dührssen & Jorswieck, 1965)**

Gruppe	Statistische Daten		
	Anzahl der Patienten	Mittelwerte der Krankenhausaufenthalte	Streuung
(1) Neurotiker im Jahrfünft vor der PSA	125	26,09	26,87
Neurotiker im Jahrfünft nach der PSA	125	5,90	14,16
Neurotiker vor der Anamnese im Jahrfünft vor 1958	100	25,55	30,40
(2) Neurotiker auf der Warteliste im Jahrfünft nach 1958	100	23,91	28,65
AOK – Versicherte im Jahrfünft vor 1958	100	10,04	18,20
AOK – Versicherte im Jahrfünft nach 1958	100	11,70	19,45

Die Krankenhaustage der Patientengruppen (1 und 2) waren vor der Therapie bzw. Anamnese signifikant zahlreicher, als bei den übrigen Versicherten. Das blieb bei den unbehandelten Patienten auch in den fünf Jahren nach der Therapie so. Die Krankenhaustage der behandelten Patienten jedoch sanken in den fünf Jahren nach Abschluss der Therapie nicht nur signifikant unter den Wert vor der Therapie, sondern auch signifikant unter den Durchschnittswert der allgemeinen AOK-Mitglieder.

Angesichts der vielfältigen inhaltlichen und methodischen Probleme der Wirksamkeitsforschung wird diese Studie als eine pragmatische Lösung angesehen, sich zunächst auf den Kosten-Nutzen-Aspekt zu beschränken. Eine derartige Analyse ist also durchaus hilfreich, um die Frage zu beantworten, ob Psychotherapie als notwendige Krankenbehandlung bezahlt werden soll.

Das erste bedeutsame Sammelreferat zur Wirksamkeit psychotherapeutischer Behandlung legten **Meltzoff & Kornreich (1970)** vor. Es wurden 101 Untersuchungen mit Kontrollgruppendesign berücksichtigt und zwischen qualitativ hochwertigen und fragwürdigen Studien unterschieden. In beiden Gruppen waren die Ergebnisse positiv, wobei die Wirksamkeit von Psychotherapie bei den Studien mit anspruchsvollen Design deutlicher nachweisbar war.

Kächele (1992) resümiert in seinem Psyche-Artikel drei Studien, in denen psychoanalytische Therapie verglichen wurde mit der Standardanalyse. Dabei zeigte sich einerseits (Hamburg et al., 1967, zit. bei Kächele, 1992), dass sich die hier erfassten Ergebnisse für ca. 10.000 Analysen nicht wesentlich voneinander unterscheiden, andererseits zeigte sich zum Beispiel, dass bei Patienten mit schwerer Symptomatik und einer neurotischen Persönlichkeit Psychoanalysen zu besseren Therapieergebnissen führten, während bei der Diagnose einer Borderline-Persönlichkeit bei der Standardanalyse deutlich mehr Verschlechterungen als bei der psychoanalytischen Psychotherapie registriert wurden (Bachrach et al., 1985).

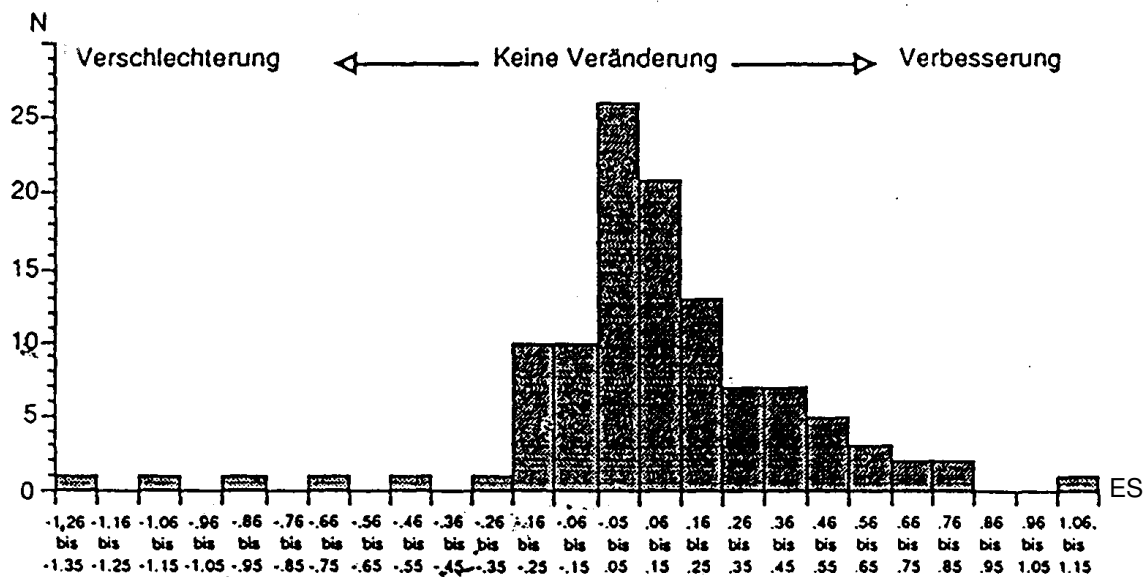
Kächele verweist dann darauf, dass manche Patienten sich unter psychoanalytischer Behandlung auch verschlechtern. Diese, auch generell für das Feld der Psychotherapie zutreffende Tatsache, wurde von **A. Bergin** 1971 in einer systematischen Übersicht über Verschlechterungseffekte erstmals für alle Psychotherapieformen dokumentiert (Bergin, 1971; Vergleich auch Strupp et al. 1977).

**Grawe (1992)** hat sich ausführlich mit der Kritik Eysencks an der Wirksamkeitsbehauptung von Psychotherapie beschäftigt. Er verweist darauf, dass die spontane Remissionsrate der Störungen typischer Psychotherapiepatienten weit geringer ist, als Eysenck angenommen hatte. **Grawe, Bernauer und Donati (1990)** zählten etwa 900 kontrollierte psychotherapeutische Outcome-Studien mit klinischen Patientenstichproben, die von den Anfängen der Psychotherapieforschung bis Anfang der 80er Jahre durchgeführt wurden.

579 Studien enthielten Kontrollgruppen. Grawe und Mitarbeiter haben nun durchschnittliche Effektstärken über alle in der Untersuchung erhobenen Veränderungsmaße berechnet. Sie sind dabei vorgegangen wie **Smith, Glass und Miller (1980)** dies zuerst für den Psychotherapiebereich getan haben: Sie haben das Ausmaß der Veränderungen in den Behandlungsgruppen in Effektstärken ausgedrückt, in dem man für jedes erhobene Maß die Mittelwerte vor und nach der Behandlung voneinander abzieht und die Differenz in das Verhältnis setzt zur Streuung der Prä-Werte.

Die entsprechenden Berechnungen konnten für 111 Kontrollgruppen vorgenommen werden. Die Verteilung der dabei resultierenden Effektstärken ist in der Abbildung 1 zu sehen.

**Abb. 1: Verteilung integrierter Effektstärken von 111 unbehandelten Kontrollgruppen  
(aus Grawe, 1992, S. 134)**



Verteilung der integrierten Effektstärken (ES), die für insgesamt 111 unbehandelte Kontrollgruppen aus kontrollierten Psychotherapiestudien gefunden wurden. Auf der Abszisse die Effektstärkenkategorien. ES um Null bedeuten keine Veränderungen, positive ES eine Verbesserung, negative ES eine Verschlechterung. Pro Kontrollbedingung wurde eine durchschnittliche ES über alle in der Kontrollbedingung erhobenen Maße berechnet. Auf der Ordinate die Häufigkeit, mit der ES der betreffenden Höhe gefunden wurden.

Es zeigt sich ein deutlicher Schwerpunkt der Effektstärken um Null und leicht darüber. Ausgeprägteren Verbesserungen stehen fast im gleichen Ausmaß Verschlechterungen gegenüber.

Nach Grawe (1992) ist dies das echte Ausmaß spontaner Remissionen, mit denen man die Wirkungen psychotherapeutischer Beeinflussungen vergleichen muss.

**Smith et al. (1980)** kommen nach Sichtung von 475 Erfolgsstudien zur Psychotherapie zu dem Ergebnis, dass bei einer durchschnittlichen Effekt-Messgröße von 0.85 sich 50 % aller behandelten Psychotherapiepatienten wohler fühlten als 75 % einer therapiebedürftigen aber unbehandelten Kontrollgruppe.

Auch **Luborsky et al. (1975)** ermittelten in einer Übersicht über 100 Erfolgsstudien einen durchgängigen Vorteil der Therapie- gegenüber der Kontrollgruppe.

Dass die Kombination von Psychopharmaka plus Psychotherapie sich einer rein psychopharmakologischen Behandlung durchgängig als überlegen erweist, zu diesem Ergebnis kommt die Übersichtsarbeit von **Rohrmeier (1982)**. Nach **Fischer und Klein (1997)** legen zahlreiche Studien die Annahme nahe, dass Psychopharmaka zwar zu rascher Symptombesserung führen können, langfristig jedoch keine stabilen positiven Resultate haben, während Psychotherapie langsamer wirkt, jedoch zu dauerhaft stabilen Änderungen führt. Nachfolgende Übersichtsarbeiten und Meta-Analysen hätten die durchschnittliche Überlegenheit von Psychotherapie gegenüber der Null-Bedingung und Psychopharmakotherapie nicht wieder in Frage gestellt.

So kann die Frage der ersten ergebnisorientierten Forschungsperiode: „Wirkt Psychotherapie, und hat sie im Durchschnitt positive Effekte, die mit anderen Mitteln nicht erreicht werden können?“ grundsätzlich positiv beantwortet werden (vgl. auch Grawe, 1992; Lamberts und Bergin, 1994).

## **2.2 Die 2. Phase der empirischen Psychotherapieforschung**

Sie ist zentriert auf den Zusammenhang von Verlauf und Ergebnis. Die Kernfrage lautet: „Welche Variablen im Prozess bewirken ein positives Ergebnis?“

Als bedeutendstes Projekt in der zweiten Forschungsphase wird die Langzeitstudie der **Menninger Klinik** in Topeka an 42 Langzeit-/Psychoanalysepatienten angesehen. Dieses zeitlich sehr aufwendige psychoanalytische Therapieforschungsprojekt erstreckte sich über 30 Jahre, von 1952 bis 1982. Die Patienten, die an die Menninger Foundation zur Behandlung kamen, umfassten das Spektrum schwerer Charakterstörungen, Suchtverhalten und Perversionen sowie narzisstische oder Borderline-Zustände, neben den schweren Formen der klassischen Psychosen.

Die Absicht des PRP (Psychotherapy Research Project) der Menninger Foundation bestand darin, die Behandlungsverläufe und so weit wie möglich auch das Leben der Patienten nach der Behandlung weiter zu verfolgen. Die Hälfte der Gruppe der 42 Patienten machten eine Psychoanalyse, die andere Hälfte eine psychoanalytische Psychotherapie. Die Zuweisung der Patienten zu den Behandlungsformen wurde klinisch entschieden, denn eine wichtige methodische Entscheidung des Menninger Projektes bestand darin, eine naturalistische Studie durchzuführen.

Das Forschungsvorhaben sollte auf die klinische Arbeit möglichst ohne Einfluss sein. Die Behandlungszeiträume erstreckten sich von ½ Jahr bis zu zwölf Jahren. Alle Patienten konnten im vorgesehenen Zwei- oder Drei-Jahresabstand für eine katamnestische Befragung erreicht werden.

Über 1/3 der Patienten konnte während eines Zeitraumes zwischen zwölf und 24 Jahren nach Behandlungsende beobachtet werden, einige Beobachtungsfristen erstreckten sich sogar bis zu 30 Jahren.

Nach **Wallerstein (1990)** bestand das Ziel dieses Forschungsprojektes darin, soviel wie möglich darüber in Erfahrung zu bringen, *welche* Veränderungen in der Psychoanalyse und in anderen psychoanalytisch fundierten Psychotherapien tatsächlich stattfinden (die Frage nach dem Ergebnis) und *wie* diese Veränderungen im Laufe der Zeit durch welche Variablen beim Patienten in der Therapie und beim Therapeuten sowie in der sich entwickelnden Lebenssituation – die alle zusammen diese Veränderungen mitbestimmen – geschehen oder herbeigeführt werden (die Frage nach dem Prozess). Wichtige Resultate dieses Forschungsprojektes sind:

- Alle Behandlungen (einschließlich der Psychoanalysen) enthielten mehr stützende Elemente als ursprünglich vorgesehen, und diese stützenden Elemente waren in erheblich größerem Umfang für die erreichten Veränderungen ausschlaggebend, als man zunächst vorausgesehen hatte.
- Der Unterschied zwischen struktureller Veränderung durch einsichtsändernde Verfahren und durch supportive Techniken ist um vieles weniger ausgeprägt, als ursprünglich angenommen wurde. Generell sind die Unterschiede von Psychoanalyse und psychoanalytischer Psychotherapie sehr gering. Kächele (1992) verweist in diesem Zusammenhang auf „die praktisch-klinische Bedeutung der Therapieforschung als eine antiideologische Produktivkraft“. Je sorgfältiger einzelne Therapieverläufe und die Ergebnisse analysiert werden, je fallnäher in der Auswertung gearbeitet wird, desto weniger fallen die Unterschiede zwischen den als standardisiert ausgegebenen Therapieformen ins Gewicht: Die Variabilität von Veränderungsprozessen tritt in den Vordergrund.
- In der Praxis existieren die verschiedenen Formen der Psychoanalyse bzw. analytischen Psychotherapie oder stützenden Psychotherapie kaum in einer auch nur annähernd idealen oder reinen Form.

Wallerstein betont, „dass

- die wirklichen Behandlungen in der tatsächlichen Praxis im Kern unentwirrbar Vermischungen von mehr oder weniger expressiv-deutenden und mehr oder weniger stützend-stabilisierenden Elementen darstellen;
- beinahe alle Behandlungen (sogar die vermeintlich reinen Psychoanalysen) viel mehr an stützenden Komponenten enthalten, als ihnen gemeinhin zugestanden wird;
- die Gesamtergebnisse, die durch die eher ‚analytischen‘ im Gegensatz zu den eher ‚stützenden‘ Behandlungen erzielt wurden, weit weniger auseinander liegen, als unsere Erwartungen im Hinblick auf diese unterschiedlichen Modalitäten ahnen lassen; und
- die Veränderungen, die mit Behandlungen an beiden Enden des Spektrums erzielt wurden, ihrer Natur und Dauerhaftigkeit nach sich ebenfalls weniger stark unterscheiden, als man erwartete, und häufig sogar kaum voneinander zu unterscheiden sind“.

Im folgenden berichten wir über drei weitere langjährige Forschungsprojekte, die in den Vereinigten Staaten die zweite Phase der psychoanalytischen Wirkungsforschung bestimmen. Wir beziehen uns dabei im wesent-

lichen auf Bachrach (1991), Fonagay et al. (1999), sowie Wallerstein (1997). Dabei wird sich zeigen, dass, obwohl manche Studien-Designs nicht mehr dem heutigen Standard entsprechen, die Resultate dennoch auch aus heutiger Sicht noch Aktualität besitzen. Das Ringen um die dem Gegenstand der Wirksamkeit psychoanalytischer Therapie adäquate Forschungsmethode hält an.

### **Das Columbia Psychoanalytic Center Research Project (Weber, Elison, Moss)**

700 Psychoanalyse-Patienten der Columbia University Psychoanalytic Clinic wurden von Analytikern in Ausbildung (Kandidaten) behandelt. Die Analysen fanden zwischen 1945 und 1971 statt.

Die Daten wurden aus Falldarstellungen, von Patienten-, Analytiker-, Supervisoren- und Instituts-Protokollen gewonnen. Die Klinikprotokolle wurden durch neun ausgebildete Analytiker nach 38 demographischen und 36 klinischen Variablen sowie neun Ich-Stärke-Skalen zu Beginn der Behandlung und bei Beendigung oder, wenn das früher der Fall war, beim Abschluss der Ausbildung des Kandidaten ausgewählt.

In einer erneuten, verbesserten Analyse der Daten (Weber et al., 1985, zit. b. Wallerstein, 1997) wurde die gesamte Stichprobe in zwei Teile aufgeteilt: 1. die Fälle, die zwischen 1945 und 1961 behandelt wurden und zu denen Daten über den Erfolg aufgrund der Beurteilungen der Protokolle durch vollständig ausgebildete Analytiker vorlagen, und 2. die Fälle, die zwischen 1962 und 1971 behandelt wurden, bei denen Informationen über das Ergebnis direkt von dem behandelnden Analytiker und den Patienten erhoben worden waren. In beiden Stichproben wurde der endgültige Erfolg der Fälle in Bezug auf die Analysierbarkeit und den therapeutischen Nutzen eingeschätzt.

Das höchste Ausmaß an therapeutischem Nutzen wurde in den Psychoanalysen festgestellt, die über den Ausbildungsabschluss des Kandidaten hinaus behandelt wurden (und damit zu den am längsten dauernden gehörten) und während eines Teilabschnitts der Analyse von den erfahrensten Kandidaten behandelt wurden. 52 % endeten in dem „gegenseitigen Einvernehmen“, dass der größtmögliche Erfolg erreicht wurde, 91 % wurden als „verbessert“ eingeschätzt, obwohl bei nur 43 % eingeschätzt wurde, dass sich ein analytischer Prozess entwickelt habe.

Dagegen zeigten Analysen, die vor dem Ende der Ausbildung des Kandidaten abgeschlossen waren (und dementsprechend die kürzesten Analysen mit gleichzeitig den relativ unerfahrensten Kandidaten waren), den geringsten Erfolg. Nur 26 % endeten mit maximalem Erfolg; in mehr als der Hälfte der Fälle wurde die Behandlung einseitig beendet: die Patienten galten als nicht gebessert oder verschlechtert.

Insgesamt gab es nur einen geringen Zusammenhang zwischen Analysierbarkeit und Therapieerfolg (Korrelationen zwischen .30 bis .40). Aber 89 % der analysierbaren Patienten beendeten ihre Therapie mit „maximalen Erfolg“, 78 % wurden als „sehr gebessert“ beurteilt, und 91 % wurden als funktionierend im neurotischen Bereich diagnostiziert. Im Gegensatz dazu beendeten nur 47 % der Patienten, die als nicht analysierbar eingeschätzt waren, die Therapie mit „maximalen Erfolg“, und nur 40 % wurden als „sehr gebessert“ eingestuft. Natürlich erreichten Patienten, die als analysierbar eingestuft wurden, einen größeren therapeutischen Erfolg als Patienten, die als solche eingeschätzt wurden, keinen psychoanalytischen Prozess einzuge-

hen. Weder die Analysierbarkeit noch der therapeutische Erfolg erwiesen sich als relativ gut vorhersehbar durch die Einschätzung des Klinikchefs (vier erfahrene Analytiker). Das galt auch für die Charakteristika der Patienten, wie sie zu Beginn der Behandlung beurteilt worden waren, und die elementaren Charakteristika der analysierenden Kandidaten oder für die Kombinationen dieser Faktoren. Sie erwiesen sich alle als schlechte Indikatoren. Die Behandlungsdauer war der einzige unabhängig gemessene Faktor, der durchgängig mit dem therapeutischen Erfolg und der Analysierbarkeit korrelierte.

Im zweiten Teil der Studie wurden 36 Analysen und 41 Psychotherapien, die von 43 Kandidaten zwischen 1962 und 1971 durchgeführt wurden, genau untersucht, weil für sie vollständige Daten vorlagen. Detaillierte Fragebögen, einschließlich Rating-Skalen, wurden an die behandelnden Kandidaten, Patienten und Supervisoren verteilt.

Daten über den Analytiker wurden durch Institutsaufzeichnungen und Ratings der Lehranalytiker erhoben. Während 96 % der analytischen Fälle nach Einschätzung ihrer Analytiker als gebessert eingestuft wurden (1/3 beendete die Therapie mit „maximalem Erfolg“, und ein anderes Viertel beendete die Behandlung in gegenseitiger Übereinstimmung, aber ohne maximalen Erfolg), wurde nur bei 50 % der analytischen Fälle die Entwicklung eines analytischen Prozesses als gegeben eingeschätzt. Die Analysierbarkeit wurde wieder in Beziehung zum therapeutischen Erfolg gesetzt. 78 % der analysierbaren Fälle wurden als „sehr gebessert“ im Gegensatz zu 29 % der Fälle eingestuft, bei denen sich ein analytischer Prozess nicht entwickelt hatte.

75 % beendeten die Therapie mit maximalem Erfolg im Gegensatz zu 38 % der unanalysierbaren Fälle. Die Dauer der Behandlung war wieder der einzige Faktor, der substantiell mit der Analysierbarkeit und dem therapeutischen Erfolg in Verbindung stand.

Zusammenfassend lässt sich sagen, dass die Begrenzungen der Studie darin zu sehen sind, dass die behandelnden Analytiker alle unerfahren waren, der Behandlungsprozess nicht unabhängig untersucht worden war, ebenso nicht der persönliche Einfluss des Analytikers auf das Ergebnis, und dass die Daten mehr deskriptiv als intrapsychisch waren. Die Verwendung von gering entlohnenden klinischen Behandlungen und die Tatsache, dass die behandelnden Analytiker auch die Hauptbeurteiler der als „gebessert“ eingestuften Fälle waren, führte zu einer systematischen Verzerrung.

### **The Boston Psychoanalytic Institute Prediction Study**

Die Studie von Sashin et al., 1975, zit. b. Wallerstein, 1997) überprüfte die Ergebnisse einer Anzahl von Patienten, die durch Kandidaten-Analytiker unter Supervision behandelt wurden. Die Patienten waren als geeignet ausgewählt worden für eine Analyse bei einem Analytiker im Kandidatenstatus, deshalb können sie nicht notwendigerweise als repräsentativ für psychoanalytische Fälle im allgemeinen gelten. Die Behandlungsergebnisse von 31 Patienten mit neurotischen Krankheiten wurden retrospektiv vom behandelnden Analytiker beurteilt, die Mehrzahl der Patienten war entweder als hysterisch, zwanghaft oder als gemischte Neurosen diagnostiziert. Die Behandlungsergebnisse wurden nach der Beendigung der Analysen auf einer globalen Veränderungsskala, einer Skala, die Lebenssituationen zum Behandlungsende prüfte, und sechs klinischen Skalen, die Symptomeinschränkung, symptombezogenes Unwohlsein, Arbeitsfähigkeit, sexuelle Anpassung, interpersonale Beziehungen und Einsicht abdeckte, beurteilt.

Ergebnis: Über 1/4 der Patienten beendeten die Behandlung vorzeitig - wenigstens vom Standpunkt des Analytikers aus gesehen -, 3/4 der behandelten Patienten wurden als gebessert beurteilt, 6 % zeigten eine bedeutsame Verschlechterung. Diejenigen, die eine längere Behandlung hatten, zeigten eher günstige Ergebnisse. Die Zustimmung des Analytikers zum Behandlungsende war mit einem guten Ergebnis verknüpft.

In weiteren Untersuchungen des Boston Psychoanalytic Institute führten **Kantrowitz et al.** (zit. bei Bachrach 1991) eine prospektive Studie mit der Frage nach der Eignung für Psychoanalyse durch, in der sie sowohl psychologische Tests als auch übliche klinische Behandlungsmethoden anwandten. Die Studie erbrachte eine Reihe von wichtigen Resultaten, so auch weitere Bestätigung über die Variationsbreite psychoanalytischer Techniken unter den Psychoanalytikern und auch die Wichtigkeit des Matchings von Patient und Analytiker als ein Prädiktor für die Langzeitwirkung.

Ein anderes wichtiges Ergebnis dieser Studie war die Beobachtung, dass die Analytiker optimistischer bzgl. des Erfolges ihrer Patienten waren als dies durch die unabhängigen psychologischen Tests nahegelegt wurde. Weiterhin zeigte diese Studie, abgesehen von den relativ kleinen Sample, dass einige Patienten noch nach der Behandlung Verbesserungen erzielten.

Bei anderen Patienten zeigte sich, dass die Verbesserungen zum Behandlungsende nicht über einen langen Zeitraum beibehalten werden konnten. Die Bewertung der Studie: der Einsatz von Kandidaten und das retrospektive Design setzen der Generalisierbarkeit Grenzen, die weiteren Studien machten jedoch deutlich, wie wichtig die Kombination Patient - Analytiker ist.

Die zweite Phase des Bostoner Projektes (Kantrowitz, 1997) diente der Evaluation der Stabilität des Ergebnisses der prospektiven Längsschnittuntersuchung bei 22 Patienten und der Faktoren, die dazu beigetragen haben. Die Forscher wollten wissen, ob das Befinden der Patienten ein Jahr nach der Behandlung ihr Befinden langfristig voraussagen lässt. 17 Patienten wurden fünf bis zehn Jahre nach Ende der Analyse mit individuellen katamnestischen Interviews nachuntersucht, und es zeigte sich, dass viele Patienten ihren Therapiefortschritt halten, wenn nicht verbessern konnten, einige sich aber auch verschlechtert hatten. Dabei stellte sich heraus, dass „kein vor der Analyse eingeschätztes Merkmal [(z.B. Krankheitsgrad), weder einzeln noch in Kombination miteinander, die Stabilität der Errungenschaften voraus“ sagte] (Kantrowitz, 1997, S. 66). Dagegen hat das Zusammenpassen von Patient und Analytiker einen bedeutsamen Einfluss auf das Ergebnis der Analyse. „Der dynamische Aspekt der Interaktionen zwischen Patient und Analytiker, ihre Resonanzen und Dissonanzen und ihre gemeinsame Fähigkeit oder Unfähigkeit in der Entwicklung der „blinden Flecke“ oder der Überbrückung der Unterschiede, die sich über den Verlauf der analytischen Arbeit entwickeln, sind wahrscheinlich zentral für das Ergebnis“ (Kantrowitz, 1997, S. 69).

#### Die Untersuchungen des New York Psychoanalytic Institute (Erle, Goldberg)

Die Untersuchungen waren naturalistisch angelegt mit einem Prä-Post-Vergleich. Die Analysen wurden von Kandidaten und erfahrenen Analytikern des New York Psychoanalytic Institute durchgeführt. Die Erfolge wurden durch die Analytiker beurteilt.



Sample: In der ersten Studie wurden 40 Behandlungen (davon 75 % Frauen), die von analytischen Kandidaten zwischen 1967 und 1969 durchgeführt wurden, untersucht. Die Mehrzahl der Patienten waren junge Erwachsene und ihre Diagnosen, getroffen durch eine Anzahl von erfahrenen Analytiker, passten in das neurotische Spektrum. In der zweiten Studie wurde eine ähnliche Anzahl von Patienten von erfahrenen Analytikern behandelt. In diesem Sample waren 60 % Männer, das Alter war breiter gestreut. Es gab einige stärker gestörten Patienten in diesem Sample, jedoch waren sie eine Minderheit.

In der ersten Studie wurden die Behandlungen durch Supervision begleitet. 2/3 der Behandlungen endeten mit gegenseitigem Einverständnis, 3/4 der Behandlungen dauerte länger als zwei Jahre. In der zweiten Studie gab es keine Behandlungs-Supervision, aber die Therapeuten waren alles erfahrene Analytiker.

Die erste Studie enthielt keine formalen Maße für die Wirksamkeit. Beurteilungen wurden auf der Basis des Eindrucks der Kandidaten über das Ausmaß des Fortschrittes gemacht.

Analytiker und Supervisoren beurteilten die Veränderung auf einer eigens entwickelten Rating-Skala. In der zweiten Studie füllte jeder behandelnde Analytiker einen halb-strukturierten Fragebogen aus, in dem Informationen zur Begründung für die psychoanalytische Behandlung, den Behandlungsprozess und eine allgemeine Beschreibung der Analysen enthalten war. Es gab keine operationalisierten Maße der Veränderung. Es wurde eine Methode, die der Ziel-Erreichungs-Skala ähnlich war, angewendet, die Behandlungsziele wurden jedoch nicht im vorhinein festgehalten. Outcome-Parameter waren Selbstachtung, Symptom-Veränderung, Veränderungen in der Abwehr, Veränderungen in Beziehungen, Veränderungen in Persönlichkeitszügen.

In der ersten Studie wurden über 80 % der Patienten mit „substantiell profitiert“ eingeschätzt, wenn sie mehr als vier Jahre in Behandlung waren, aber nur 12 % wurden mit dieser Kategorie bewertet, wenn sie weniger als zwei Jahre Analyse hatten. Über alles gesehen zeigte die Mehrheit eine gewisse Behandlungswirkung, aber 40 % wurden als wenig verändert eingeschätzt.

In der zweiten Studie waren es nur 55 %, bei denen ein therapeutischer Nutzen festgestellt worden war. Aber in allen Fällen, die sich verändert hatten, war am Anfang die Analysierbarkeit festgestellt worden. Nicht alle Fälle, die als analysierbar am Anfang eingeschätzt worden waren, zeigten eine bedeutsame Veränderung. In beiden Studien gab es eine starke Beziehung zwischen Länge der Behandlung und Outcome.

Bewertung der Studien: Diese Studien sind prinzipiell von historischen Interessen, weil die Outcome-Ratings von Klinikern als nicht-valide betrachtet werden. Ein besonders interessanter Aspekt der Studie ist die Beziehung zwischen der Etablierung eines psychoanalytischen Prozesses (beurteilt durch den Therapeut) und dem Outcome (Fonagay et al., 1999).

In allen vier referierten Projekten ging es darum, tatsächlich stattfindende oder stattgefundene psychoanalytische/psychotherapeutische Behandlungen auf ihre Wirksamkeit bzw. die Faktoren zu untersuchen, die Einfluss auf die Wirksamkeit haben. Dabei sollte der analytische Prozess, wenn er zeitgleich mit der Untersuchung ablief, durch die Untersuchung soweit wie möglich ungestört bleiben. Man wollte den psychoanalytischen Behandlungsprozess möglichst unverzerrt studieren können.

Zu einer Entfernung von der klinischen Praxis führt dagegen die **experimentelle Gruppenvergleichsmethodik** der 2. Phase der empirischen Psychotherapieforschung.

Riedel und Schneider-Düker (1991) setzten sich mit den Kontextbedingungen einer Metaanalyse für den deutschsprachigen Raum von **Wittmann und Matt** (1986) auseinander. In 58 % der insgesamt 86 Behandlungsvergleiche handelte es sich um sog. Analogstudien, das heißt, die „Therapien“ waren nur auf Grund der Studien zustande gekommen. Ohne das Angebot hätten die Versuchspersonen keine Behandlung aufgesucht. Die externe Validität solcher Studien ist im Rahmen der Psychotherapie-Forschung höchst zweifelhaft (Heckrath und Dohmen, 1994). Die Behandlungsdauer betrug zwischen einer und 60 Therapiesitzungen: Bei Psychoanalyse im Durchschnitt 13,8 Stunden, bei klientenzentrierter Therapie 10,8 Stunden, bei Verhaltenstherapie 15,2 Stunden und bei Entspannung und Suggestion 15,8 Stunden. Es ist offensichtlich, dass viele Patienten eine zu geringe Anzahl von Therapiesitzungen hatten.

Durch das Kriterium „Vorhandensein einer Kontrollgruppe“ zur Aufnahme in die Metaanalyse waren Studien ausgeschlossen mit Therapien von mehr als 20 Sitzungen, Patienten mit schweren Störungen und psychoanalytische Behandlungsmethoden, die alle keine Kontrollgruppe enthielten.

Riedel und Schneider-Düker (1991, S. 26) charakterisieren die für eine Metaanalyse geeignete Studie wie folgt:

- „Behandelt wird vorwiegend im universitären Beratungssetting oder in der Schule durch Psychologen, Studenten und Lehrer; Ärzte kommen nicht vor.
- Behandelt werden eher milde Symptomatiken bei Klienten, die häufig für die Untersuchung rekrutiert wurden.
- 80 % erhalten weniger als 20 Sitzungen.
- Verhaltenstherapeutische Methoden dominieren, gefolgt von Gesprächspsychotherapie und Entspannungsverfahren.
- Die Untersuchung bzw. die Therapie ist als Forschungsstudie konzipiert.“

Antworten auf die Fragestellung der Phase 2 versuchen auch Untersuchungen zu geben, die einen Vergleich zwischen verschiedenen Therapieschulen erlauben.

Klassisches Beispiel für diese Forschungsrichtung ist die Untersuchung von **Sloane et al.** (1975): Die Autoren verglichen Verhaltenstherapie und Psychoanalyse (psychoanalytische Kurztherapie bei einer durchschnittlichen Behandlungsdauer von 13-14 Sitzungen). Die Therapien wurden, im Unterschied zu einer Reihe anderer Untersuchungen, von erfahrenen Therapeuten durchgeführt. Mit sehr genauen Meßmethoden und Kontrollgruppen wurden die Wirkungen und Unterschiede dieser Therapien erfasst. Die von den Verhaltenstherapeuten eingesetzten Techniken umfassten alle Komponenten verhaltenstherapeutischer Behandlungsmöglichkeiten.

Beide Therapieformen stellten sich als gleichermaßen wirksam heraus, gleichzeitig waren die Ergebnisse deutlich besser als die einer Kontrollgruppe (Warteliste). Dabei war man überrascht festzustellen, dass die Therapeuten sich während der Therapie anders verhielten als man es aufgrund ihrer Zugehörigkeit zu ihrer

Therapieschule vermutet hätte. Die Verhaltenstherapeuten waren insgesamt aktiver, erteilten Ratschläge und Instruktionen und gaben ebenso viele Interpretationen wie die Analytiker.

(Dass die Analytiker sich nicht immer schulengerecht wie Psychoanalytiker verhalten, hatte ja auch Wallerstein s.o. festgestellt, z.B. indem sie, falls notwendig, Ich–stabilisierend und supportiv vorgehen.)

Weitere Untersuchungen mit dem Ziel, die beste Therapieform für bestimmte Krankheitsbilder zu finden:

**Grawe** (1976) und **Plog** (1976), zitiert nach Bozok & Bühler (1988): Sie verglichen Verhaltenstherapie (VT) mit Gesprächspsychotherapie (GT) bei Agoraphobikern. Die Autoren kamen zu der Auffassung, dass sich beide Therapieformen nicht im wesentlichen dadurch unterscheiden, dass eine Therapieform erfolgreicher war als die andere, sondern dadurch, dass sie unterschiedliche Effekte bei unterschiedlichen Patienten bewirken. So konnten sie feststellen, dass VT besonders erfolgreich bei Patienten mit starken phobischen Ängsten und einem ausgeprägten Leidensdruck wirkte, während bei solchen Patienten, bei denen der subjektive Leidensdruck nicht so weitgehend auf die phobische Symptomatik konzentriert war, die phobischen Symptome nicht reduziert wurden. VT scheint sich spezifischer auf die phobische Symptomatik auszuwirken als GT.

Die Studie von **Rosenthal** (zitiert bei Bozok und Bühler, 1988), bei der mehr als 100 verhaltenstherapeutische Studien überprüft worden waren, zeigte, dass sehr häufig Placebo-Bedingungen, die zur Kontrolle eingesetzt wurden, gleich gute oder nur wenig schwächere Erfolge zeigten als die untersuchten Techniken.

Es wurde daraus geschlossen, dass eine verhaltenstherapeutische Technik einer Placebo-Kondition immer dann nicht überlegen war, wenn diese den Patienten gleichermaßen glaubwürdig erschien wie die verhaltenstherapeutische Technik.

Singer und Luborsky (1975, S.1006) ziehen in ihrer Metaanalyse zum Wirksamkeitsvergleich verschiedener Therapieformen die häufig zitierte Schlussfolgerung: „Everybody has won and all must have prizes.“ Zahlreiche Studien zum Schulenvergleich konnten bis heute dieses sog. Vogel-Dodo-Verdikt aus „Alice im Wunderland“ nicht schlüssig widerlegen.

Auch **Donald F. Klein** (zit. nach Heinzl et. al.,1996) kommt nach einer gründlichen Revision der bisherigen Effektivitätsforschung zu dem Resümee, dass gesicherte Erkenntnisse über spezifische Wirkungen einzelner Therapieformen einschließlich Kosten-Nutzen-Analysen zwar sehr wichtig wären für individuelle und soziale Entscheidungen, dass aber bislang alle Versuche, von verschiedenen Therapieverfahren, Therapeutenvariablen, ja sogar Erfahrungsgraden der Therapeuten auf signifikante Unterschiede in Grad oder Dauer von Therapieeffekten zu schließen, praktisch gescheitert sind (Klein & Rabkin, 1993). In diesen Zusammenhang passt die Untersuchung von **Strupp u. Hadley**, die nachweisen konnten, dass nicht nur nicht die Therapeutenvariablen für den Therapieeffekt entscheidend sind, sondern ganz allgemein menschliche Eigenschaften. Strupp und Hadley (1979) haben versucht, den spezifischen Wirkanteil technischer Interventionsmittel („technical skills“) gegen den allgemeinen Wirkanteil einer guten Beziehung („qualities inherent in any good human relationship“) abzuwägen. Sie verglichen die therapeutischen Ergebnisse einer Gruppe von professionellen

Therapeuten mit vieljähriger beruflicher Erfahrung mit den von Lientherapeuten ohne psychotherapeutische Ausbildung.

In die Gruppe der Lientherapeuten wurden College-Professoren aufgenommen, von denen bekannt war, dass sie im Umgang mit Studenten hilfreiche und verständnisvolle Beziehungen aufnehmen konnten. Insgesamt 64 Patienten waren Studenten mit Auffälligkeiten in den MMPI-Skalen Depression, Angst und Introversion. Die Ergebnisse zeigten, dass die College-Professoren in der Therapeutenrolle im Wesentlichen vergleichbare Erfolge wie die Gruppen der erfahrenen Therapeuten bei einem vergleichbaren „Patientengut“ erzielen konnten.

Das professionelle Geschick der Psychotherapeuten konnte im Vergleich zum alltagspsychologischen Geschick der Laien – unter der Voraussetzung einer tragfähigen Beziehung in beiden Therapiegruppen – keinen eigenständigen „Therapiezugewinn“ für sich beanspruchen. Das Ergebnis überrascht, weil die Lientherapeuten hauptsächlich auf solch unprofessionelle Interventionsmittel wie „informelle Konversation“ und „Ratgeben“ zurückgegriffen haben.

Psychotherapie ist dann erfolgreich, wenn der Patient bereit und motiviert ist, das Setting anzunehmen und daran zu partizipieren. Strupp und Hadley (1979) schlussfolgern aus ihren Ergebnissen, dass die positiven Veränderungen in ihrer Studie im Wesentlichen auf die gute zwischenmenschliche Beziehung zurückzuführen ist. „Therapeutische Veränderungen scheinen sich dort einzustellen, wo sich ein Zusammenspiel ergibt zwischen einem Patienten, der sich auf eine solche Beziehung einlassen kann (der z.B. nicht zu großen Widerstand zeigt und hochmotiviert ist) und einem Therapeuten, in dessen Interventionen die Patienten Fürsorge und echtes Interesse erfahren“ (S.1135 f.). Sowohl in der Gruppe der professionellen Therapeuten als auch in der Gruppe der Lientherapeuten gab es effiziente und wenig effiziente Therapeut-Patient-Paare.

Der Unterschied zwischen „erfolgreich“ und „weniger erfolgreich“ scheint demzufolge auch mit dem therapeutischen Geschick und dem „Zusammenpassen“ von Therapeut und Patient zu tun zu haben.

Es gilt deshalb weiterhin als Tatsache, dass die bisher untersuchten Therapieverfahren trotz offensichtlich sehr unterschiedlichem Therapeutenverhalten im Durchschnitt zu annähernd äquivalenten Effekten führen. Diese Äquivalenz der Therapiewirkung kann auf sogenannte „unspezifische Wirkfaktoren“ (common factors) zurückgeführt werden. Diese sind in jeder hilfreichen zwischenmenschlichen Beziehung wirksam und werden damit in allen erfolgreichen Psychotherapien verwirklicht.

Meyer (1990) macht für fast alle Therapieerfolge im Wesentlichen unspezifische bzw. „kommunale“ Faktoren verantwortlich.

Auch **Orlinsky und Howard** (1986) sehen im Zentrum des therapeutischen Prozesses die therapeutische Beziehung. Die Qualität der Beziehung ist sowohl vom Verhalten des Therapeuten als auch von dem des Patienten abhängig. Drei Dimensionen der Therapeut-Patient-Beziehung sind von Bedeutung:

- Engagement (real investment) des Therapeuten/des Patienten, Glaubwürdigkeit/Sicherheit des Therapeuten, Motivation des Patienten, Echtheit des Therapeuten;
- empathische Resonanz (empathatic resonance): Einstellung des Therapeuten/Patienten, Merkmale des

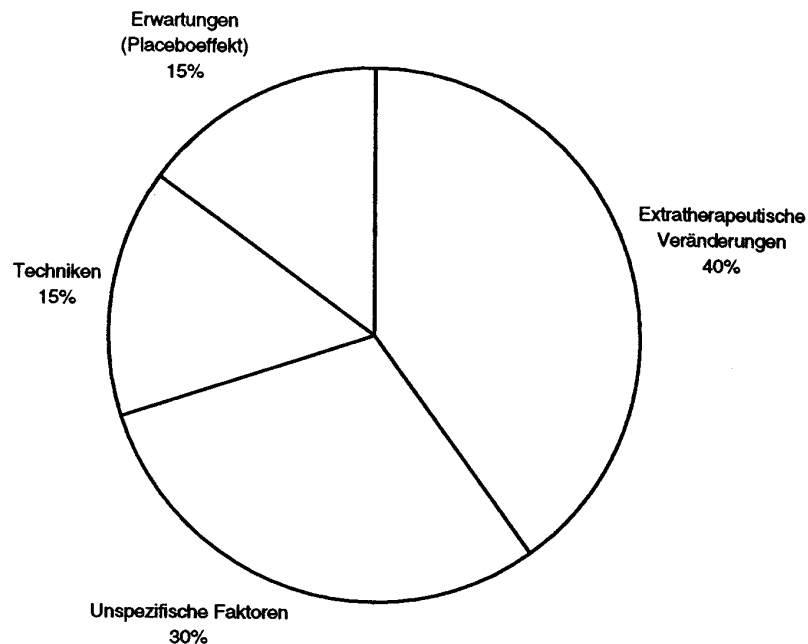
Ausdrucks beim Therapeuten/Patienten (Stimme, Sprache, Mimik), gegenseitige Resonanz;  
– gegenseitiges Vertrauen (mutual affirmation): Wärme, Akzeptanz, Vertrauen des Therapeuten/Patienten, reziprokes Vertrauen.

Grundlage von Veränderungen ist eine tragfähige Therapeut–Patient–Beziehung. Diese Erkenntnis ist durch eine Vielzahl auch empirischer Untersuchungen (Übersicht bei Orlinsky und Howard 1986) abgesichert. Diese Autoren erwarten positive Auswirkungen einer tragfähigen therapeutischen Beziehung sowohl auf die unmittelbaren Ergebnisse einzelner Therapiesitzungen, als auch langfristige Veränderungen des Verhaltens des Patienten außerhalb des therapeutischen Rahmens.

Schließlich betont auch Grawe (1994), dass für die Einzeltherapie die Bedeutung der Qualität der Therapiebeziehung über alle Zweifel erhaben nachgewiesen ist, und zwar für ganz unterschiedliche Therapieformen. Sie kann als der empirisch am besten abgesicherte Wirkfaktor der Psychotherapie angesehen werden.

Es erscheint damit für die breite Mehrheit der Therapieforscher klar, dass die Therapieschulen-spezifischen Techniken und Variablen der Therapiegestaltung nicht der (allein) ausschlaggebende Faktor sind. **Lambert** (1992) zeigt denn auch, basierend auf einer Übersicht der Outcome-Forschung (Abb. 2), wie sich die therapeutischen Faktoren hinsichtlich ihres Beitrages zur Verbesserung von Psychotherapiepatienten verteilen.

**Abb. 2: Verbesserung bei Psychotherapiepatienten als eine Funktion therapeutischer Faktoren (aus Lambert, 1992, S. 97)**



- Extratherapeutische Veränderungen (40 %): Diese Faktoren sind ein Teil des Klienten (z.B. Ich-Stärke) und Teil seiner Umgebung (wie zufällige Ereignisse, soziale Unterstützung). Diese Faktoren helfen bei der Genesung, unabhängig von der Teilnahme an einer Therapie.
- Erwartungen (Placeboeffekt) (15 %): Der Anteil an der Verbesserung, der aus dem Wissen des Klienten resultiert, dass er/sie behandelt wird und von der Glaubwürdigkeit bestimmter Behandlungstechniken und Grundprinzipien.
- Techniken (15 %): Einzigartige Faktoren spezifischer Therapien (wie Biofeedback, Hypnose oder systematische Desensibilisierung).
- Unspezifische Faktoren (30 %): Hier sind eine Reihe von Variablen eingeschlossen, die in einer Vielzahl von Therapien vorgefunden werden, unabhängig von der theoretischen Orientierung des Therapeuten: z.B. Empathie, Wärme, Akzeptanz, Ermutigung ein Risiko einzugehen, usw.

Im Unterschied zu Lambert (1992) sind Butler und Strupp (1986) der Ansicht, dass es sehr schwer ist, die spezifischen Einflussgrößen systematisch von den unspezifischen zu trennen, weil die Wechselwirkungen zwischen ihnen eng miteinander verflochten sind.

Die Autoren kommen zu dem Schluss, dass die Dichotomisierung zwischen spezifischen und unspezifischen Faktoren unberechtigt und unproduktiv sei, weil die eigentliche therapeutische Situation der interpersonelle Kontext ist. In diesem Sinne vermuten Bozok und Bühler (1988, S.126) „unspezifische Faktoren seien das ‚Material‘, die spezifischen Faktoren, die ‚Form‘. So wie Form und Material untrennbar sind, so muss jede unspezifische Wirkung in einer bestimmten Gestalt erscheinen“.

### **2.3 Kritik an Effektstudien und experimentellem Gruppenvergleich**

Das Problem der Gültigkeit der Forschungsergebnisse und die Frage nach der „richtigen“ Methode:

- Bei der Frage nach der Gültigkeit der Ergebnisse der vergleichenden Effektstudien für die Praxis tauchen Probleme auf. Ist es überhaupt statthaft ist, so unterschiedlich angelegte Studien mit ganz unterschiedlichen Zielsetzungen und Erfolgsdefinitionen miteinander zu vergleichen, sie also praktisch alle über einen Kamm zu scheren und ist das Resultat von Forschung nicht durch Zuverlässigkeit und Gültigkeit des Messinstrumentes beeinflusst?
- In Bezug auf den experimentellen Gruppenvergleich: Werden hier überhaupt relevante Ergebnisse produziert, da ja , wie wir gesehen haben, in der Praxis des Analytikers oder Verhaltenstherapeuten möglicherweise etwas ganz anderes passiert als in den schulenspezifischen ex- oder impliziten Behandlungsvorschriften vorgesehen. So stünde hier ja womöglich eine von Praktikern nicht praktizierte Theorie auf dem Prüfstand von Wirkung und nicht das, was wirklich in der Sitzung vom Therapeuten/Analytiker praktiziert wird.

**Lambert und Hill** (1994, S.74 ff.) weisen nach einer ausführlichen Revision der verschiedenen Messinstrumente für Therapieeffekte unter anderem darauf hin, dass

- es schwer sei, die Ziele von Therapie adäquat zu gewichten,
- die Einschätzung von Verhaltensänderungen durch Therapie unausweichlich vermischt sei mit Erwartungen,
- Maßeinheiten zur Veränderungsmessung, die von individuell zugeschnittenen Therapiezielen stammen, kaum vergleichbar seien,
- die Art des Messinstrumentes, Fragebogens, usw. einen Einfluss auf die Einschätzungen hat, die gemessenen Therapieeffekte unterschiedlich ausfallen, je nach dem, aus welcher Quelle sie stammen.

Zum letzten Aspekt hat sich gezeigt, dass die Ergebnisse aus den fünf untersuchten Quellen (Selbsteinschätzung, trainierter Beobachter, wichtige andere Person, Therapeut, instrumentelle Messung) nicht übereinstimmen (Vergleich Kazdin, 1994, S. 40). Diese und viele andere gravierende Unterschiede in den Effektstudien lassen nach Meinung der Autoren die Methode der Meta-Analyse sehr fragwürdig erscheinen. Viele der bisherigen Therapiestudien scheinen wegen der genannten Mängel vermutlich nur begrenzt aussagekräftig.

In einem Übersichtsreferat über zahlreiche Sammelstudien und Meta-Analysen zählen **Bozok und Bühler** (1988) eine Reihe wichtiger Kritikpunkte an diesen Studien auf:

- Uniformitätsmythos (Kiesler, 1966, zit. bei Bozok u. Bühler, 1988): Hierbei handelt es sich um die irrige Annahme, dass bei den untersuchten Patienten- und Therapeutengruppen um uniforme Personen handele, die man ohne Einschränkungen miteinander vergleichen könnte.
- Zu kleine Stichprobe: Viele Studien beruhen auf zu geringen Patientenzahlen, so dass statistische Aussagen und Generalisierungen problematisch sind. Nach Bergin und Garfield (1994) ist die durchschnittliche Stichprobengröße bei allen ihnen verfügbaren Untersuchungen zwölf.
- Die Patientenauswahl entspricht nicht den praktischen Gegebenheiten (freiwillige Versuchspersonen, „Patienten“ mit überwiegend leichten Störungen).
- Die Therapeuten sind nicht erfahren genug: Nach Parloff (1982) waren die meisten Therapeuten Doktoranden mit einer mittleren therapeutischen Erfahrung von drei Jahren.
- Die Erfolgskriterien sind nicht adäquat bzw. uneinheitlich.
- Uneinheitliche Länge der Therapien: Nach Henry et al. (1994) ist die durchschnittliche Gesamtzahl der Sitzungen der untersuchten Therapien zwischen 20 und 150, die Mehrheit unter 50.
- Nachfolgeuntersuchungen fehlen.

Fischer und Klein (1997) kommen zu dem Urteil, dass die Validität der in experimentellen Studien gewonnenen Ergebnisse fraglich ist. Für den Einzelfall in der klinischen Praxis erscheinen die Ergebnisse wenig relevant. Für die gegenwärtige Forschungsphase sei deshalb die Wendung zur realen klinischen Praxis kennzeichnend.

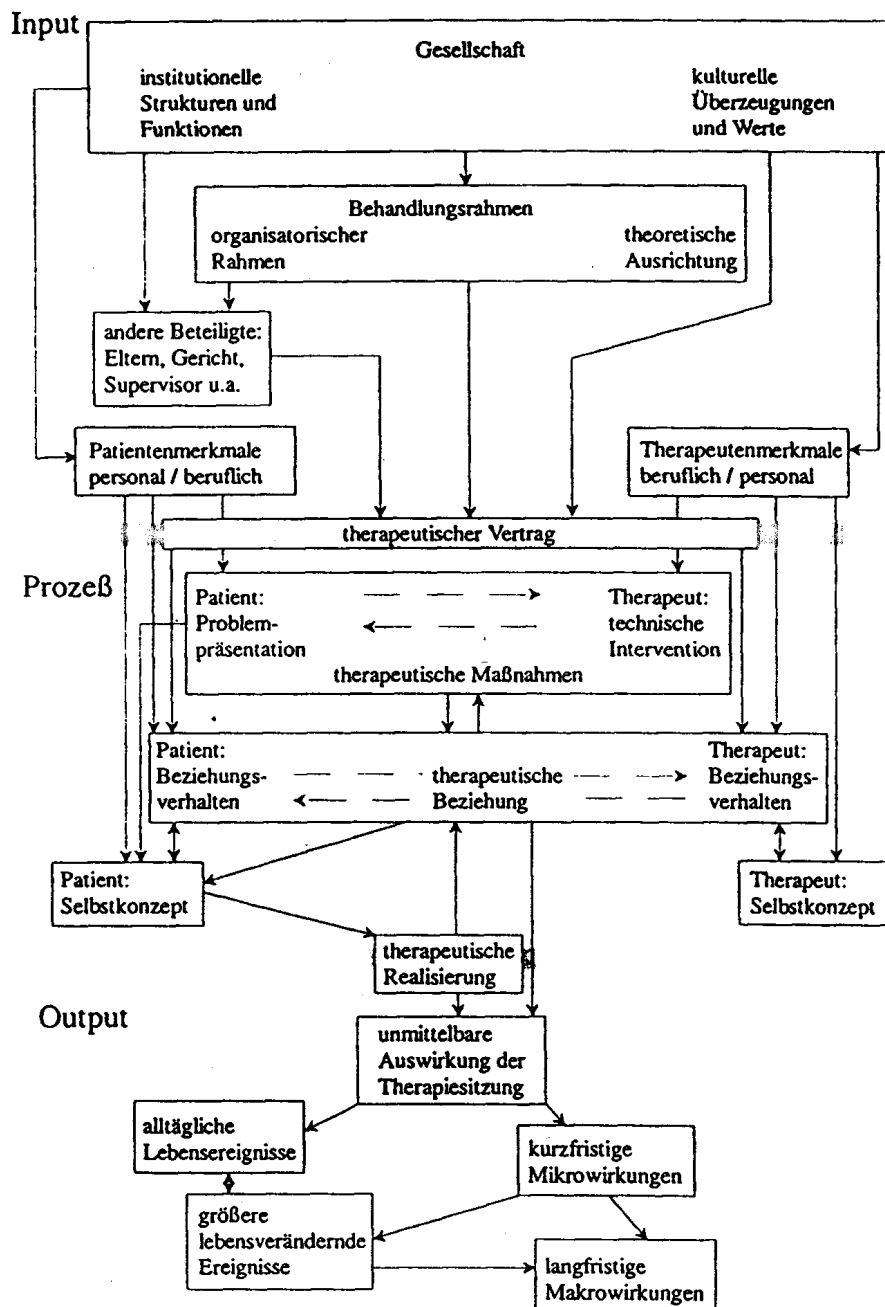
## **2.4 Die 3. Phase der empirischen Psychotherapieforschung**

„Die 3. Phase der Psychotherapieforschung wendet sich wieder der naturalistischen Untersuchungsperspektive zu, hält jedoch daran fest, dass prozessuale Faktoren kontrolliert in eine Studie einzubeziehen sind.“ (Kächele, 1992; S. 271)

Zur Verdeutlichung der ungeheuren Vielschichtigkeit und Komplexität des Therapieprozesses eignet sich ein Modell von **Orlinski und Howard** (1986) (s. Abb. 2a).



Abb 2a: Modell des Therapieprozesses  
(aus Orlinski und Howard, 1986)



In diesem Modell werden Input-, Output- und Prozessvariablen unterschieden. Zu den Inputvariablen gehören Charakteristika des Patienten und des Therapeuten sowie die institutionellen, theoretischen und gesellschaftlichen Rahmenbedingungen. Auswirkungen und Konsequenzen von Psychotherapie werden in den Output-Variablen zusammengefasst.

Zu den Prozessvariablen zählen alle Aspekte der Zusammenarbeit von Therapeut und Patient. Es werden folgende Kategorien unterschieden:

- therapeutischer Kontrakt
- therapeutische Interventionen
- therapeutische Beziehung
- Selbstkonzepte des Therapeuten/ des Patienten
- therapeutische Realisierung

Die Metaanalyse von über 1.100 Forschungsergebnissen (Orlinski und Howard, 1986) ergab, dass eine große Zahl untersuchter Merkmale (z.B. Therapiedauer, Einsatz bestimmter Interventionstechniken, Aspekte der Therapeut–Patient–Beziehung) mit den Therapieerfolgen in Zusammenhang stehen. Das Problem der gegenwärtigen Prozessforschung besteht darin, dass sich quantitative Analysen, die verallgemeinerbare Aussagen über den Zusammenhang von Prozessmerkmalen mit Erfolgsmaßen erlauben, sich auf wenige, häufig isolierte Variablen beschränken (Groß und Riedel, 1995). In diesem Zusammenhang betonen Fischer und Klein die Notwendigkeit, von einer variablenorientierten zur fallorientierten Perspektive überzugehen. Die Frage in Phase drei der Forschungsgeschichte ist demnach nicht mehr “Was soll geschehen?” sondern die eher mikrologische Frage “Was geschieht wirklich?” Es scheint, dass der Kommunikationsprozess zwischen Therapeut und Patient den Therapieerfolg in hohem Maße beeinflusst, was in Einzelfallstudien untersucht werden kann.

Neben dem Problem der Generalisierbarkeit der Ergebnisse aus Einzelfallstudien besteht aber auch die Forderung nach methodisch exakter Erfassung dessen, was zwischen den wichtigen Bezugspersonen und dem Patienten geschah und geschieht und wie sich die Beziehung zwischen Patient und Therapeut konstituiert. So steht die Untersuchung der immer wieder neu inszenierten Beziehungsmuster im Mittelpunkt vielfältiger Forschungsbemühungen.

In den heutigen Methoden empirischer Psychotherapieforschung stehen typische, sich wiederholende, maladaptive Muster (Zyklen) zwischenmenschlichen Verhaltens im Vordergrund, z.B. das zentrale Beziehungs-Konflikt-Thema (CCRT, Luborsky), die SASB-geleitete rekonstruktive Therapie (SASB-RLT, Benjamin, s. Kap. 11.2.3.2), die Erfahrungen des Patienten mit seiner Therapeutenbeziehung (PERT, Hoffmann und Gill) u.a.

Es würde den Rahmen unserer Arbeit sprengen, die Methoden hier jetzt im einzelnen vorzustellen.

„Während die Forschungsmethodik von Phase zwei meist mit apriorischen Setzungen arbeitete, z.B. die Forderung nach homogenen Patientengruppen durch diagnostische Kategorien verwirklicht sah, ist die kliniknahe Methodik eher aufdeckend und suchend (heuristisch) orientiert“ (Fischer und Klein, 1997, S. 26). Es gehe darum, in individuellen Fallkonstellationen strukturelle Gemeinsamkeiten im Material aufzusuchen und nachzuweisen.

Zwei Forschungslinien bestimmen die Psychotherapieforschung (wir folgen hier weitgehend der Darstellung von **Hermann Faller**, in Faller und Frommer, 1994), zum einen die kasuistische Forschung mit der auf Breuer und Freud zurückgehenden Tradition und der aus der psychologischen Psychotherapieforschung stammende Ansatz, psychotherapeutische Prozesse quantifizierend abzubilden. Goudsmit (1989) schreibt (übersetzt von uns): „Es gibt eine tiefe Trennung zwischen zwei Traditionen in der Psychologie. Man kann sie nennen, die ‚Draußen-tradition‘ (outside tradition) und die ‚Innen-tradition‘ (inside tradition). Die folgenden Begriffe beschreiben diese Traditionen:

Zum einen: positivistisch, reduktionistisch, erklärend, experimentell, nomothetisch, quantitativ, objektiv, ‚hart‘. Zum anderen: phänomenologisch, ganzheitlich, verstehend, beschreibend, idiographisch, qualitativ, subjektiv, ‚weich‘.“

**Mertens** (1994, Kap.14) stellt die hier gemeinten ‚zwei Welten‘ unter den Begriffen ‚empirischer Objektivismus‘ und ‚Tiefenhermeneutik‘ einander gegenüber.

Beide Ansätze haben nach Mertens ihre Möglichkeiten und Begrenzungen. Die Vorteile der kasuistischen Methode bestünden darin, die Subjektivität des Individuums in ihrer ganzen Reichhaltigkeit, Breite und Tiefe abzubilden. Empirische Forschung eröffne dagegen die Möglichkeit, operationalisierbare Merkmale psychotherapeutischer Behandlung in intersubjektiv vergleichbarer Weise zu erfassen und einzelfallüberschreitend Aussagen zu Ursache-Wirkungs-Zusammenhängen machen zu können. Dass beide Forschungsrichtungen sich nicht gegeneinander ausschließen müssen, sondern ihrer beider Vorteile miteinander kombinierbar sind, zeigt **Leuzinger-Bohleber**, 1994 in ihren aggregierten Einzelfallstudien, die auf einer „mittleren Abstraktionsebene“ einerseits der Idiosynkrasie des Einzelfalles gerecht werden, andererseits aber auch mit Hilfe statistischer Verfahren Gemeinsamkeiten und Unterschiede zeigen.

Nachteil der Fallgeschichte ist, so Mertens, der subjektive Einfluss des Forschers, der für den Leser nicht deutlich erkennbar und bewertbar ist. Um als ernst zu nehmender Forschungsbeitrag in der psychotherapeutischen und wissenschaftlichen Öffentlichkeit zur Kenntnis genommen zu werden, ist eine den Regeln der Beweiskraft und Logik folgende Argumentation vorzulegen.

Der Fallgeschichte fehlt wegen der subjektiven Einfärbung Beweiskraft. Faller (1994, S. 17) bezieht sich auf Spence (1989), der eine Forschungsmethodik fordert, die es dem Leser erlaubt, angenommene Verknüpfungen zwischen Tatsachen und Schlussfolgerungen einzuschätzen und ihm Möglichkeiten des Widerspruchs, der Widerlegung und Falsifikation eröffnen. „In dieser neuen Art der Darstellung hat das Material das Wort – nicht der Therapeut.“ (Spence)

Die quantifizierende Forschung kann nur dasjenige erfassen, was objektivierend messbar ist. Dies birgt die Gefahr in sich, dem Gegenstand „Psychotherapie“ nicht gerecht zu werden. Hieraus erklären sich die für den

Praktiker uninteressanten Resultate und die fehlende Rezeption von methodisch sehr aufwendig durchgeführten Untersuchungen.

**Qualitative Forschung** kann eine Brücke zwischen den beiden Polen schlagen und die Vorteile beider Richtungen, den Reichtum des Subjektiven und die Stringenz der Forschungslogik integrieren. Qualitative Forschung kann, so verstanden, als Grundlagenforschung angesehen werden (Vergleich Grawe, 1988; Tress, 1989, Kächele, 1992). Es geht hier um die Beschreibung, Interpretation und Erklärung dessen, „was der Fall ist“ (Vergleich Tress, 1989).

*Allgemeine Prinzipien qualitativer Forschung* sind (nach Wilke in Faller/Frommer, 1994, S. 78 f.):

- Das Prinzip der Offenheit: unvoreingenommener Blick auf das Material
- Verzicht auf Hypothesenbildung ex ante: Kategorienbildung nur aus dem zu untersuchenden Material selbst, und nicht aus Daten anderer Studien
- Der Prozesscharakter der Forschung: Theoriebildung als Ergebnis der Wechselbeziehung des Forschers mit seinen Daten
- Das Prinzip der Explikation und Begründung: Offenlegung und Nachvollziehbarkeit eines jeden Schrittes der Analyse
- Forderung nach Flexibilität: Fortwährende mögliche Überprüfung von Vorannahmen, Hypothesen und Ergebnissen
- Die Sicht der Subjekte: Versuch des Forschers, den Prozess subjektiver Interpretationen und Sinnherstellungen nachzuvollziehen
- Der Prozess der Beschreibung, Rekonstruktion und Entdeckung: Herausarbeiten inhärenter Strukturen und Regelmäßigkeiten des Materials
- Wahrheit und Objektivität: Kriterien, wie Handlungs- und kommunikative Validierung werden angelegt

Welche Studien sind nun bislang in der Phase drei der Psychotherapieforschung von besonderer Bedeutung? Wir werden sehen, dass dies mehrheitlich Studien mit den überkommenen Fragestellungen der ersten und zweiten Phase der Psychotherapieforschung sind. Der Anspruch, prozessuale Faktoren kontrolliert in einer Studie zu untersuchen, rückte schon bald wegen des durch Grawe - s.u. Punkt 3) - ausgelösten neuen Rechtfertigungsdrucks in den Hintergrund.

1) In den frühen 80er Jahren wurde in Berlin eine großangelegte Studie konzipiert und Ende der 80er durchgeführt (Rudolf et al., 1988b), die als die **Berliner Psychotherapiestudie** bekannt wurde.

Sie ist eine Beobachtungsstudie im Feld analytischer Psychotherapie von Berliner Patienten. Im Mittelpunkt des Interesses steht das Zustandekommen von Psychotherapien, die stufenweise Realisierung des therapeutischen Kontaktes vom diagnostischen Erstgespräch über die Therapieplanung bis zur Durchführung der Behandlung. Die Erfassung und Bewertung der therapeutischen Arbeitsbeziehung ist ein besonderes Anliegen der Studie. Es wird der diagnostisch-therapeutische Alltag von analytischen Psychotherapeuten unter

naturalistischen Bedingungen wissenschaftlich begleitet. Die wichtigsten Fragestellungen der Studie lassen sich wie folgt zusammenfassen:

- Welche Patienten wenden sich an welche Institution?
- Welches sind die Prädiktoren für die Indikation zur Psychotherapie, ihre Realisierung und ihren erfolgreichen Abschluss?
- Welche Formen therapeutischer Arbeitsbeziehungen entwickeln sich unter welchen Bedingungen?
- Welchen Anteil hat die Persönlichkeit des Psychotherapeuten/Psychoanalytikers an diesem Geschehen?

Das Miteinander von Patienten und Therapeut im Spiegel ihrer Selbst- und Fremdeinschätzung steht im Mittelpunkt (Rudolf et al., 1988b).

Die zahlreichen in der Studie verwendeten, z. T. neuentwickelten Erhebungsinstrumente sollen intrapsychische und interpersonale Befunde auf unterschiedlichen Ebenen der Wahrnehmung dokumentieren. Es sind dies die folgenden sieben Ebenen: 1. Ebene des Expertenurteils (Therapeut sieht Patienten), 2. Ebene der therapeutischen Beziehung (Therapeut sieht therapeutische Zusammenarbeit), 3. Ebene der Gegenübertragung (Therapeut erlebt seine Reaktion auf den Patienten), 4. Ebene der Selbsteinschätzung (Therapeut sieht sich selbst), 5. Ebene der Selbsteinschätzung (Patient sieht sich selbst), 6. Ebene der therapeutischen Beziehung (Patient sieht therapeutische Zusammenarbeit), 7. Ebene der Übertragung (Patient sieht Therapeut). Die insgesamt 348 ambulanten und stationären Behandlungen stammten aus 17 psychoanalytischen Praxen (DGPT-Psychoanalytiker) und sieben klinischen Institutionen. Die Stundenfrequenz der Analysen betrug 2 - 3 Wochenstunden, die der dynamischen Psychotherapien waren sehr variabel. Der Untersuchungszeitraum betrug insgesamt 3 1/2 Jahre. Die katamnestische Nachuntersuchung, 3 1/2 Jahre nach der Erstuntersuchung, zu der ca. 50% der Psychoanalysen beendet waren, schloss die Studie ab.

#### *Die verwendeten Ergebnismaße:*

- Globalmaß (Direkte Einschätzungen der Veränderungen): Einschätzung des Ausmaßes des Symptomwegs falls (durch die Therapeuten und Patienten) und der Umstrukturierung der Persönlichkeit (durch die Therapeuten).
- Indirekte Veränderungsmessung durch Vorher-Nachher-Vergleich: Für die Patienten-Selbsteinschätzung anhand von elf Skalen und die Therapeuten-Einschätzung in zehn Skalen wird ein Vorher-Nachher-Vergleich vorgenommen. Die zehn bzw. elf Skalen werden zur Auswertung zu jeweils drei Gruppen zusammengefasst, die im Schwerpunkt somatische, psychische und Beziehungsstörungen beinhalten. Als Besserung wird definiert, dass in den drei zusammengefassten Befundskalen mindestens 50 % der anfangs als schwer oder mittelschwer eingeschätzten Auffälligkeiten abschließend nur als leicht oder als fehlend markiert werden.
- Kritische Differenz. Als Kriterium für die Besserung im Vorher-Nachher-Vergleich der Skalen wird die kriti-

sche Differenz verwendet. Sie bestimmt sich aus der Skalenreliabilität (Cronbachs Alpha) und beträgt im Durchschnitt 1 - 1,5 Standardabweichungen.

Rudolf et al. (1994) untersuchten 44 psychoanalytisch behandelte Patienten im Hinblick auf den qualitativen und quantitativen Therapieerfolg. Die Ergebnisse wurden verglichen mit 56 dynamischen Psychotherapien und 164 stationären Behandlungen. Der Vergleich der Symptome, Diagnosen und Motivation der Patienten in den unterschiedlichen Therapiesettings vor der Therapie zeigt, dass sehr unterschiedliche Patientengruppen in die verschiedenen Therapien gelangen. Deshalb scheint Randomisierung nicht geeignet, Gruppen in verschiedenen therapeutischen Settings zu vergleichen.

*Ergebnisse:* mit Hilfe unterschiedlicher Ergebniskriterien konnte gezeigt werden, dass die Psychoanalysepatienten sich ausgesprochen stark verbesserten und im Vergleich der drei Therapien die jeweils besten Behandlungsergebnisse erzielen. Verschiedene Wirkungsmaße und verschiedene Perspektiven (Patienten/Therapeuten) zeigten weitere interessante Unterschiede. Patienten berichteten vorrangig über Verbesserungen bei somatischen, ängstlichen und depressiven Beschwerden. Therapeuten dagegen berichteten über deutliche Verbesserungen bei Beziehungseigenschaften und dem Verhalten. Die relativ häufigsten und relativ ausgeprägtesten Befundveränderungen im Vorher-Nachher-Vergleich bei psychoanalytischen Therapien betreffen aus der Sicht der Patienten und der Therapeuten die Bereiche:

- körpernahe Angstsymptomatik (z.B. plötzliche Angstzustände, Atemnot),
- Depression (z.B. Selbstmordgedanken),
- ängstlich im Kontakt (z.B. leicht verunsichert, gekränkt),
- körperliche Symptomklage (z.B. Nackenschmerzen, Kopfschmerzen).

Im Rahmen der Berliner Psychotherapiestudie wurde auch die initiale Situation zwischen 739 therapiesuchenden Patienten und ihren Therapeuten daraufhin untersucht, welche prognostischen Faktoren sie im Hinblick auf Indikationsentscheidung und Behandlungsverlauf enthält (Rudolf et al., 1988a).

Die Suche nach den Prädiktoren des Therapiegeschehens erfolgt dadurch, dass Zusammenhänge zwischen den diagnostischen Anfangsbefunden und den Therapieverlaufskriterien überprüft wurden. Als zusammenfassendes Ergebnis lässt sich festhalten, dass die diagnostische Initialsituation eine Vielzahl von Merkmalen enthält, die einen gesicherten statistischen Zusammenhang mit den Verlaufskriterien aufweisen und deshalb als prognostische Prädiktoren betrachtet werden können.

Es konnten vier Prädikatorengruppen identifiziert werden:

*Erste Prädikatorengruppe:* Therapeutische Zusammenarbeit und persönliche Wertschätzung

Diese Prädikatorengruppe ist die bei weitem bedeutsamste. Die Merkmale beziehen sich nicht auf den Patienten als behandlungsbedürftigen Kranken, sondern auf seine Eignung und Bereitschaft zur therapeutischen Zusammenarbeit mit einem gleichfalls geeigneten und motivierten Therapeuten vor dem Hintergrund wechselseitiger persönlicher Wertschätzung. Es ist der Respekt und das wohlwollende Interesse für die

Persönlichkeit des Patienten. Die Wertschätzung kann sowohl auf spontaner zwischenmenschlicher Sympathie als auch auf professionell geschulter Verständigung beruhen.

*Zweite Prädikatorengruppe: Soziale Kompetenz und soziale Belastung*

Der Patient erlebt sich vorwiegend als körperlich krank und leitet daraus eine ausgeprägte Rücksichtsforderung ab. Er signalisiert vielfältige Vermeidungsstrategien. Der Patient betont die Macht und Stärke des Therapeuten und die eigene Ohnmacht und Schonungsbedürftigkeit.

*Dritte Prädikatorengruppe: Emotionaler Kontakt*

Es geht um die Schwierigkeit des Patienten im Kontakt, speziell um seine Balance von Annäherung und Rückzug, Nähe und Abgrenzung. Das interaktionelle Klima ist gefärbt durch misstrauische Kontaktvermeidung, durch kühle Distanz oder auch Enttäuschungsärger und Vorwurfshaltung.

*Vierte Prädikatorengruppe: Stützende Introjekte*

Diese Gruppe von Prädikatoren liefert das Bild eines Selbst, das sich auf positive Introjekte stützt und sich auf die biographische Erfahrung von eigener durchsetzungsfähiger Aktivität und Verantwortlichkeit bezieht. Es wird hier nicht die kranke und hilfsbedürftige Seite des Patienten betont, sondern eher seine Ich-starke und kontaktfreudige Seite hervorgehoben.

Ein weiterer Befund macht die Bedeutung des Therapeuten als prognostischen Experten für den Verlauf der psychotherapeutischen Behandlung deutlich. Das Therapeutenurteil zu Behandlungsbeginn ist ein guter Prädiktor für Verlauf und Ergebnis der Behandlung.

Im Rahmen der Berliner Psychotherapiestudie wurden auch Muster therapeutischer Zusammenarbeit und ihre Beziehung zum Therapieergebnis (Grande et al., 1988) untersucht. Grundlage bildeten 123 dokumentierte Therapieverläufe im Bereich der stationären Psychotherapie. Über insgesamt fünf Messungen zur therapeutischen Arbeitsbeziehung (von Seiten des Therapeuten und von Seiten des Patienten) wurden verschiedene Clusteranalysen berechnet, um bestimmte Typen des Verlaufs der therapeutischen Zusammenarbeit zu identifizieren. Vier Cluster wurden ausgewählt.

Im Cluster 1 findet sich eine sehr günstige Einschätzung der Arbeitsbeziehung durch Therapeut und Patient. Die Patienten dieses Clusters sind weniger krank als die anderen Patientengruppen, sie sind gut motiviert, kontaktbereit und entwicklungsfähig. Ihre Einstellungen und Fähigkeiten sind für die therapeutische Zusammenarbeit förderlich und die therapeutische Beziehung ist von wechselseitiger Wertschätzung getragen.

In Cluster 2 wird die therapeutische Arbeitsbeziehung von beiden Seiten auf mittleren Niveau eingeschätzt. Diese Patienten scheinen ihre Problematik selbst recht gut zu erkennen, im Hinblick auf den therapeutischen Kontakt sind sie jedoch ängstlich und zurückhaltend. Sie beurteilen den Therapeuten zunächst eher vorsichtig und nehmen die psychotherapeutische Behandlung eher aus Gefügigkeit auf sich. Die Patienten haben

eine geringere Ausbildung, weshalb ihnen die Verbalisierung emotionaler Konflikte und Themen eher schwer fallen dürfte.

Im 3. Cluster finden sich gleichbleibend sehr positive Beurteilungen der Arbeitsbeziehung durch den Patienten, der Therapeut schätzt dagegen die Beziehung pessimistisch ein.

Die gefühlshafte Reaktion der Therapeuten auf die Patienten ist ungünstig, sie werden eher unsympathisch, uninteressiert und unfähig eingeschätzt. Demgegenüber betonen die Patienten ihre Bereitschaft zur therapeutischen Zusammenarbeit, sie hätten ihre Probleme schon selbst erkannt, bräuchten jedoch Hilfe, um sie zu lösen. Es wird angenommen, dass die Patienten dieses Clusters, die sehr stark mit Suchtproblemen zu tun haben, mit Hilfe ihrer betont optimistischen Beurteilungen ihre persönlichen Schwierigkeiten überspielen.

Das 4. Cluster zeigt aus der Sicht des Therapeuten eine sehr ungünstige Einschätzung der Arbeitsbeziehung mit negativer Entwicklungstendenz. Sie erscheinen am Ende der Behandlung relativ stark beeinträchtigt (Narzissmus, Angstsymptomatik u.a.). Prognostisch erscheinen sie als außerordentlich ungünstig, ihre Motivation ist schwach ausgeprägt und ihre Haltung regressiv getönt. Die Therapeuten sehen bei diesen Patienten die fehlende Bereitschaft als ausgeprägte Abwehrhaltung, wodurch der Eindruck entsteht, dass diese Patienten nicht therapeutisch kooperieren wollen oder es aufgrund ihrer Persönlichkeitsproblematik nicht können.

Die beschriebenen Beziehungscluster erlauben eine Vorhersage des Therapieergebnisses: auf der Basis der Beurteilung der Therapeuten hinsichtlich der Veränderungen des globalen Neurosebefundes und der körperlichen Symptomatik sowie der globalen strukturellen Veränderungen zeigt sich eine Rangordnung der vier Patientengruppen, die genau den Beurteilungen der therapeutischen Arbeitsbeziehung durch den Therapeuten im gesamten Verlauf der Behandlung entsprechen. Die Patienten des günstig charakterisierten ersten Clusters sind am erfolgreichsten, die Patienten des vierten Clusters zeigen dagegen vergleichsweise wenig Veränderungen.

Die zu Anfang entstandene Beziehung, wie sie vom Therapeuten wahrgenommen wird, gibt die Möglichkeiten der weiteren therapeutischen Zusammenarbeit wieder und ist daher für den Verlauf prognostisch aussagekräftig. Weiterhin konnte herausgefunden werden, dass Veränderungsmaße auf Basis von Patientenangaben insgesamt schlechter vorhergesagt werden können als Therapeutenkriterien, und dass außerdem die Zusammenhänge zwischen Arbeitsbeziehung und Therapieergebnis dann am höchsten sind, wenn die Beurteilungen der Arbeitsbeziehung und die Beurteilungen des Ergebnisses aus derselben Quelle stammen (d. h. von Patient bzw. Therapeut).

Dass die therapeutische Arbeitsbeziehung (erfasst durch die Skala TAB, das heißt durch Therapeuteneinschätzung) verlässliche Voraussagen für das Behandlungsergebnis liefert, bestätigt auch die Untersuchung von Rudolf und Manz (1993). Im Vergleich zur Therapeuteneinschätzung spielt die Patienteneinschätzung eine vergleichsweise untergeordnete Rolle. Das subjektive Befinden des Patienten bei Behandlungsende und sein subjektives Erleben der therapeutischen Zusammenarbeit haben offenbar einen geringeren Zusammen-



hang. Nach den Befunden von Rudolf et al. (1988a) sind die Aussagen des Patienten durch Abwehrhaltungen verformt. Die Untersuchung bestätigt den Therapeuten als Experten, der in der Lage ist, die Beziehungs- und Arbeitsmöglichkeiten diagnostisch-prognostisch im Behandlungsverlauf und im Hinblick auf das Behandlungsergebnis einzuschätzen.

## **2) Das Heidelberger Katamneseprojekt (v. Rad, Senf, Bräutigam, 1998)**

Im Heidelberger Katamneseprojekt wurden alle Behandlungen über einen bestimmten Zeitraum, die in der Psychosomatischen Universitätsklinik Heidelberg durchgeführt wurden, erfasst. Dies waren stationär eingeleitete und ambulant fortgeführte Einzel- und Gruppentherapien sowie primär ambulante dynamische Psychotherapien und Psychoanalysen. Vor der Behandlung wurden für jeden einzelnen Patienten drei bis fünf individuelle Therapieziele festgelegt, die dann nach Therapieende und zum Katamnesezeitpunkt (durchschnittlich 3,5 Jahre) von einem unabhängigen Untersucher überprüft wurden. Insgesamt wurden 208 Patienten untersucht, die nach ihrer diagnostischen Zuordnung sowie nach der Behandlungsform ausgewertet wurden. In Bezug auf die Symptomatik, die individuellen Therapieziele, die testpsychologischen Untersuchungen und die Zufriedenheit der Patienten ergaben sich insgesamt gute, teilweise sehr gute Ergebnisse, die fast ausnahmslos über den langen Katamnesezeitraum stabil blieben. Allerdings konnte die Gruppe der Psychoanalysepatienten hinsichtlich der Symptomatik das gute Ergebnis bei Therapieende über die Katamnesezeit hinweg nicht halten (was die Autoren irritierend und überraschend finden und sich u. a. dadurch zu erklären versuchen, dass in Psychoanalysen möglicherweise dem Selbstverwirklichungsaspekt gegenüber dem Symptom-Beseitigungsauftrag des Patienten - bewusst oder unbewusst - eine gewisse Priorität eingeräumt wurde oder wird). Dagegen schätzten fast 3/4 der Psychoanalysepatienten ihre individuellen Therapieziele als erreicht ein, im Gegensatz zur Hälfte der Patienten, die in dynamischer Psychotherapie waren. Psychosomatisch Kranke erzielten zum Katamnesezeitpunkt sehr gute Ergebnisse, insbesondere, wenn die Behandlung stationär eingeleitet worden war.

Inwieweit die individuellen Therapieziele der Psychoanalysepatienten erreicht wurden, sollte auch über ein Experten-Rating, das auf der Inhaltsanalyse von Katamneseinterviews basiert, festgestellt werden (Seibuechler-Engel et al., 1996).

Eine textanalytische Untersuchung von 53 Katamneseinterviews machte deutlich, dass zwei Jahre nach der Behandlung ein hoher Prozentsatz der Psychoanalysepatienten ihre individuellen Therapieziele als erreicht ansehen. Die Ergebnisse zeigen „gute“ oder „sehr gute“ Wirkung bei 55 % der Patienten, die in Psychoanalyse und psychoanalytischer Psychotherapie waren. Ein noch besseres Ergebnis für die Psychoanalyse zeigt das Experten-Rating der prospektiv festgelegten individuellen Therapieziele anhand der Katamneseinterviews: 72 % der psychoanalytisch behandelten gegenüber 56 % der psychotherapeutisch behandelten Patienten erreichten „gute“ oder „sehr gute“ Besserung.

Ein Bericht aus dem Heidelberger Katamneseprojekt aus dem Jahre 1989 zeigt eine interessante Dosis-Wirkungs-Beziehung (Kordy et al., 1989). Die Studie untersuchte die Abhängigkeit von Länge und Wirkung psychotherapeutischer Behandlung, indem sie Wirksamkeitsdaten der psychoanalytisch orientierten Gruppen- und Einzeltherapie verwendete. 209 Patienten mit psychoneurotischen und psychosomatischen Krank-

heiten, für die Langzeitbehandlung geplant war, wurden auf die Beziehung zwischen Dauer und Wirkung der Behandlung analysiert. Es wurden Daten von den Therapeuten (Ratings über die Symptomverbesserung, Zielerreichungsskala) und von den Patienten (Giessen-Test u.a.) über sechs Wochen nach der letzten Therapiesitzung gesammelt. Die Behandlung dauerte im Durchschnitt 2,6 Jahre und 146 Sitzungen. Die Resultate zeigen, dass die Beziehung zwischen therapeutischen Bemühungen (Länge der Behandlung und Anzahl der Sitzungen) und Therapieeffekt als ein Dosis-Wirkungsmodell beschrieben werden kann. Dementsprechend ist eine Behandlungsdauer von zweieinhalb Jahren mit 160 Sitzungen am effektivsten.

Es gibt eine schwache Tendenz bei den psychosomatischen Patienten, dass die Behandlungsdauer bis zu dreieinhalb Jahren einen Anstieg der Wirkung zur Folge hat.

3) Viel Aufsehen in Fachkreisen aber auch in der Öffentlichkeit hat die Arbeit von **Grawe und Mitarbeiter** (1994) erzeugt, weil sie zu ganz anderen Ergebnissen als denen der Phase zwei („everyone has won“) kommen. Durch ihre Metaanalyse zu Evaluation und Vergleich verschiedener Therapieformen und ihren daraus gezogenen Schlussfolgerungen lösten sie einen neuen Boom an Psychotherapieforschung aus, und zwar auch zu Fragen, die längst schon geklärt schienen, wie denen nach der Wirksamkeit von psychoanalytischer Behandlung. Grawe und Mitarbeiter ziehen aus ihren Befunden nämlich folgendes Resumé: „Die tatsächliche Ergebnislage könnte daher nicht eindeutiger sein, als sie ist. Kognitiv-behaviorale Therapie ist im Durchschnitt hochsignifikant wirksamer als psychoanalytische Therapie und Gesprächstherapie.“ (Grawe 1994, S. 670)

Die Studie ist schon bald auf erhebliche Kritik gestoßen. So vertritt Mertens (1994) in seiner Erwiderung die Auffassung, dass Effizienzvergleiche zwischen verschiedenen psychotherapeutischen Schulen mit in vieler Hinsicht unterschiedlichen Konzepten nicht mehr zeitgemäß sind. Er spricht sich dafür aus, bescheidenere, zugleich aber differenziertere Fragestellungen zu bearbeiten. Mertens wendet sich auch dagegen, die experimentell statistische Forschung als die einzige Methode der Wahrheitsfindung anzusehen.

Des weiteren wird die Tatsache, dass die meisten der zum Vergleich herangezogenen Studien zur psychoanalytischen Therapie nicht von erfahrenen psychodynamischen Therapeuten und auch nicht in einem den Standards der dynamischen Psychotherapie entsprechenden Setting durchgeführt wurden, kritisch angemerkt (Tschuschke et al., 1994).

Auch Heinzl et al. (1996, S. 7) äußern sich kritisch zum Grawe'schen Forschungsparadigma, das der Eigenart der Psychoanalyse und der analytischen Psychotherapie nicht gerecht werde. Die Psychoanalyse habe ihrem Wesen entsprechend einen qualitativen, auf das Individuum bezogenen Anspruch, der einem einzelnen Menschen in seiner persönlichen Situation und mit seinen unvergleichbaren psychischen Gegebenheiten gerecht werden wolle. Im Gegensatz dazu sei das erklärte Ziel der kognitiv-behavioralen Therapieformen, sichtbare Symptomreduktionen und Verhaltensänderungen zu erzeugen, erfassbar mit quantifizierbaren Variablen und möglichst objektiven Messmethoden in kontrollierten Studien mit prüfstatistisch verwertbaren Daten. Die von außen und auch von Grawe an die Psychotherapie herangetragenen Forderungen nach Objektivität, Transparenz, Effektivität usw. passe somit nahtlos in das Paradigma des empirischen Ob-

jektivismus und begünstige auf diese Weise die Verhaltenstherapie gegenüber der Psychoanalyse oder allen psychodynamischen Therapien in Wirkungsuntersuchungen.

4) Mit Erleichterung und Genugtuung aufseiten der Vertreter von Langzeitbehandlungen wurde nach dem „Grawe-Schock“ eine Studie aufgenommen, die gänzlich anders konzipiert war als bisherige Untersuchungen im Sektor Psychotherapie: der **„consumer-report“**, der eine demoskopische Fallstudie darstellt. Sie ist eine sozialwissenschaftliche, landesweite Verbraucherbefragung von Psychotherapie-Patienten in den USA über ihre Erfahrungen mit laufenden oder abgeschlossenen Psychotherapien. Diese Studie ließ Seligmann, 1995, zu der Auffassung gelangen (S. 267), dass Wirksamkeitsstudien die falsche Methode für die empirische Bestätigung von Psychotherapie, wie sie tatsächlich geschieht, sind, weil sie zu viele wesentliche Elemente von dem, was in der Praxis passiert, weglassen. Die im „consumer-report“, November 1995, publizierte Studie versuchte im Gegensatz dazu ein realistisches Bild von ehemaligen Psychotherapie-Patienten zu gewinnen. Ihre Bedeutung für die Psychotherapieforschung erhält sie dadurch, dass sie von Daten ausgeht, die aus tatsächlich durchgeführten, typischen psychotherapeutischen Behandlungen stammen.

7.000 Personen antworteten auf Fragen zur Behandlung psychischer Probleme sowie zur Dauer und Frequenz der Behandlung, Medikation, Art und Ausmaß der Störungen und Probleme, Kosten, Deckung durch Versicherungen, psychischer Zustand vor und nach der Behandlung u.ä. 4.100 von diesen gaben irgendeine Kombination von Behandlungen bei einem professionellen Berater oder Psychotherapeuten („mental-health-professional“), einem praktischen Arzt oder Hausarzt („family-doctor“) und Teilnahme an einer Selbsthilfegruppe an.

Diagnostisch handelte es sich vorwiegend um klassische Störungen, wie Depressionen, Angststörungen, Panikattacken im Zusammenhang mit Konfliktsituationen und Stress, wie Eheprobleme, Probleme mit Kindern, Arbeit, Verlust von nahen Personen, Alkohol, Problemen mit Drogen- und Medikamentenmissbrauch. 34 % hatten schwere psychische Beeinträchtigungen, wurden mit ihrem Leben kaum fertig.

Ausgewertet wurde mit einem kombinierten Messwert für die Therapieeffektivität, der sich zusammensetzte aus spezifischen Verbesserungen hinsichtlich jener Symptome und Probleme, die sie in die Psychotherapie führten (Anlass-Symptome); Zufriedenheit mit der Behandlung durch die Psychotherapeutinnen und Psychotherapeuten; globale Besserung hinsichtlich des allgemeinen psychischen und emotionalen Zustandes zum Zeitpunkt der Befragung, verglichen mit dem Beginn der Psychotherapie.

Ergebnisse:

- Die Behandlung durch professionelle Therapeuten führte bei den allermeisten zu einer deutlichen Besserung. Diese Ergebnisse stimmen überein mit den bisherigen Studien zur Effektivität von Psychotherapie (siehe Lambert und Bergin, 1994; Grawe, 1992).
- Die Überlegenheit einer spezifischen Therapiemethode konnte nicht nachgewiesen werden. Auch dieses Ergebnis steht in Übereinstimmung mit vorliegenden Forschungsergebnissen (z.B. Luborsky et al., 1975). Zu anderen Ergebnissen kommen jedoch Grawe et al., 1994 (s.o.).
- Längerfristige Psychotherapien führten zu einer deutlicheren Linderung von psychischen Leiden als Kurz-

therapien. Eine Behandlungsdauer von länger als zwei Jahren brachte nicht nur die besten Ergebnisse, sondern auch die größte Steigerung von Umfang und Qualität des Behandlungserfolges. Nicht nur die Besserung der Anlasssymptome fiel durch Langzeittherapie deutlicher aus, sondern es kam auch zu einer deutlichen Verbesserung der allgemeinen psychischen Funktionsfähigkeit in vielfältiger Weise.

- Professionelle Psychotherapeuten behandeln effektiver. Die Wirksamkeit der Behandlung bei Eheberatern oder Praktischen Ärzten war deutlich geringer.
- Aktive Patienten besserten sich stärker. Ihre Aktivität zeigte sich z.B. darin, dass sie sich nach Behandlungsmethoden und Erfahrungen bei früheren Patienten erkundigten, Erstgespräche bei verschiedenen Psychotherapeuten führten oder nach Erklärungen und unklaren Begriffen bzw. Diagnosen fragten o.ä.
- Psychotherapeuten nehmen Änderungen der Standardmethode hinsichtlich Dauer, Setting und Variation der Techniken vor, um der Einzigartigkeit ihrer Patienten gerecht zu werden.
- Für den psychotherapeutischen Erfolg sind, so Hutterer (1996 S. 5) in seinem Kommentar zu der Studie von Seligman, „selbstregulierende Faktoren (z.B. durch sensible Anpassung und ‚Einstellung‘ einer Technik auf die spezielle Situation der Patienten und Kombination mit anderen Techniken) besonders wirksam“. Von Bedeutung scheint auch zu sein, dass Patienten den zu ihnen subjektiv passenden Therapeuten finden. Dies erweist sich als offensichtlich erfolversprechender als eine Zuweisung auf Grund einer Indikationsstellung für eine spezifische Methode.

5) Heinzel, Breyer u. Klein legten 1996 eine **katamnestische Evaluationsstudie über ambulante Psychoanalyse in Deutschland** vor. Von ihrem Ansatz her passt sie nicht zu mikroanalytischen Fokus der dritten Forschungsphase, sondern ist eine Wirksamkeitsstudie, die an die Studie von Duehrssen und Jorswieck(1965) erinnert. Sie stellen unter anderem folgende Überlegungen an: Aus der Übersicht von Lambert und Hill (1994) gehe klar hervor, dass schon die Wahl der Meßmethoden einer Fahrt zwischen Skylla u. Charybdis gleiche. Je mehr Variablen in einem Fragebogen erfasst würden, desto umständlicher und aufwendiger sei die Untersuchung (was sich zum Beispiel bei einer Beschränkung auf Freiwillige in der Rücklaufquote niederschlägt), aber je weniger Variablen untersucht würden, desto schlechter sei es möglich ihre Ergebnisse zu generalisieren. Die Autoren kommen denn auch zu dem Schluss:

„Die Komplexität menschlichen Verhaltens und die Komplexität der Theorien und Konzepte menschlichen Verhaltens rufen äußerste Schwierigkeiten hervor, die Veränderungen, die als ein Ergebnis von Psychotherapie auftreten, zu operationalisieren“ (Lambert und Hill 1994, S. 77, Übersetzung von uns).

Für die Outcome-Forschung bevorzugen die Autoren deshalb die Alternative, nur mit anonymen Befragungen rein katamnestisch zu arbeiten, um damit die Schwierigkeit zu umgehen, die sich mit der Erforschung der multiplen Interaktionen und Wechselwirkungen zwischen Patienten und Therapeut ergeben.

Sie halten es für sinnvoller, wegen des Komplexitätsgrades des Untersuchungsgegenstandes die Messungen auf die Erfassung globaler Parameter zu beschränken. Die Autoren fassen Konzeption und Resultate der Studie wie folgt zusammen (S. 49): „Die ehemaligen Patienten einer Zufallsstichprobe aus allen DGPT- und

DGIP-Mitgliedern, die im Zeitraum Anfang 1990 bis Ende 1994 ihre analytische Therapie abgeschlossen hatten, wurden in einer anonymen katamnestischen Untersuchung um die subjektiven Einschätzungen ihres Befindens zu Beginn und am Ende der Therapie sowie zum befragten Zeitpunkt gebeten.

Weiterhin sollten sie die Häufigkeiten von Medikamentenverbrauch, Arztbesuchen, Krankschreibungstage und Krankenhausaufenthalte angeben, jeweils auf das Jahr von den drei genannten Zeitpunkten bezogen. Die Rücklaufquote betrug 65 %.

Aus den subjektiven Einschätzungen der ehemaligen Patienten sind deutliche Verbesserungen des Befindens zu erkennen, die sich nach Abschluss der Therapie nicht abschwächen, sondern zum Teil noch verstärken.

Darüber hinaus werden nennenswerte Rückgänge bei der Inanspruchnahme sonstiger Gesundheitsleistungen berichtet, und zwar sowohl bei den Arztbesuchen als auch bei Krankenhausaufenthalten. Besonders bemerkenswert ist der Rückgang der Krankschreibungstage um fast 2/3. Der Therapieerfolg ist um so größer, je schlechter der selbst-eingeschätzte Gesundheitszustand des Patienten zu Beginn der Therapie war. Ansonsten steigt er i.d.R. mit der Zahl der Sitzungen und ist bei jüngeren Patienten größer als bei älteren. Signifikante Unterschiede im Therapieerfolg zwischen den Grundberufen oder analytischen Grundorientierungen der Therapeuten sowie zwischen Einzel- und Gruppentherapie konnten nicht festgestellt werden.

Auf Grund des naturalistischen Designs und der großen Stichprobe sowie der hohen Rücklaufquote kann man davon ausgehen, dass diese Ergebnisse und die daraus abgeleiteten Aussagen für die derzeit in Deutschland ausgeübte psychoanalytische Richtlinien-Therapie repräsentativ sind."

## **6) Das Stockholmer Projekt**

Die Untersuchung von Sandell et al. (1999 u. 2001) aus Schweden hat in der psychoanalytischen community insbesondere deshalb positiv für Aufsehen gesorgt, weil hier (endlich) die Wirksamkeit von Langzeitanalysen mit der Methodik der empirischen Sozialforschung nachgewiesen werden konnte.

Die Untersuchung ist als eine Kombination aus Quer- und Längsschnittuntersuchung angelegt, und zwar so, dass empirische Daten sowohl von in Behandlung befindlichen wie auch auf Wartelisten befindlichen Patienten im Verlauf von drei Jahren zu unterschiedlichen Untersuchungszeitpunkten erhoben wurden. Die Patienten in gleicher Behandlungsphase wurden zu statistischen Untersuchungsgruppen zusammengefasst, wodurch es möglich war, Informationen über alle Behandlungsstadien zu gewinnen, nämlich über „wartende Patienten in einer Kontrollgruppe“, über Patienten in den verschiedenen Therapiephasen sowie über die Wirkung der Therapien bis zu drei Jahren nach Ende der Behandlung. An einer ursprünglichen Stichprobe von 756 Patienten bei 316 approbierten Psychotherapeuten und Psychoanalytikern wurden die folgenden Datenpools erhoben: 1. sozio-demografische und psychiatrisch relevante Daten der Patienten, 2. die objektiven Kontakte der Patienten mit dem Gesundheitssystem für den Zeitraum 1988-1996, 3. Daten, die mittels des Well-being Questionnaire (WBQ), das drei andere standardisierte Testverfahren integriert, erhoben wurden, wobei der WBQ den Patienten in den Jahren 1994-1996 zu drei mit jährlichem Abstand auseinanderliegenden Erhebungszeitpunkten mit der Bitte um Beantwortung und Rücksendung zugeschickt wurden, 4. objektiv und reliabel geratete Daten zur Erfassung der Informationen aus qualitativen Interviews, die an einer Teil-

stichprobe von 60 Patienten durchgeführt wurden, 5 soziodemographische und therapierelevante „Meinungsdaten“ der 316 Therapeuten einschließlich normierter und standardisierter Vergleichsdaten.

Der WBQ besteht aus 3 Messinstrumenten, einer halbstrukturierten Befragung zur sozialen Beziehungsgestaltung (SAS), einer Liste von 90 verschiedenen aktuellen somatischen und psychischen Beschwerden (SCL-90) und einer Selbsteinschätzungsskala hinsichtlich aktueller Befindlichkeiten (SOCS). Untersucht wurden ausschließlich Langzeitpsychotherapien mit zwei bis drei Sitzungen pro Woche und Psychoanalysen mit vier bis fünf Sitzungen pro Woche. Die Untersuchung zeichnet sich durch besonders selbstkritische Prüf- und Wirksamkeitskontrollen aus: So wurden z.B. „50 möglicherweise nicht ganz unabhängige Variablen (z.B. demographische Daten von Patienten und Therapeuten)“ (Sandell et al., 2001, S. 285) auf ihre Konstanz bzw. Veränderung hin beobachtet, um die klinischen Veränderungen bei den Patienten eindeutig auf die abhängige Variable „Dauer der Behandlungszeit“ zurückführen zu können. Merkmalsveränderungen wurden z.B. daraufhin überprüft, ob sie mit der Zeit korrelieren, also einen bloßen Zeiteffekt und keinen Behandlungseffekt darstellen.

Mit Hilfe der Datenpools konnten statistische Gruppen gebildet und abhängige und unabhängige Variablen definiert werden, deren Zusammenhänge mit statistischen Prüfverfahren untersucht wurden.

Die Auswertung der Daten erfolgte gemäß den bisherigen Veröffentlichungen hauptsächlich mit Hilfe von Mittelwertvergleichen und - wie die Autoren sagen - „die in der Psychotherapieforschung gebräuchliche standardisierte Effektstärke“. (Sandell et al., 2001, S. 285)

Die zentralen Ergebnisse der Studie bestehen darin, dass

- Langzeitpsychotherapien überhaupt nachweisbare Therapieeffekte haben,
- diese Effekte bei höherfrequenten und längeren Therapien in ihrer Langzeitwirkung über das Therapieende nachhaltiger sind als bei weniger hochfrequenten und kürzeren Therapien,
- erfahrenere Therapeuten auch die besseren Therapieeffekte erzielen,
- Therapeutinnen scheinbar besser therapieren als die männlichen Kollegen.

Es ist anzumerken, dass die Untersuchungsgruppe nicht mit einer experimentiellen Kontrollgruppe verglichen wurde, was vielleicht ein Nachteil dieser Studie ist. Man hätte dann die im Verlauf der Psychotherapien feststellbaren Veränderungen auf eine direkte empirische Basis stellen können. Mit „Kontrollgruppe“ ist in dieser Untersuchung die Eich- und Normalstichprobengruppe gemeint, die man zum Vergleich der klinischen Ausprägungen der wartenden Patienten mit den klinischen Ausprägungen der Allgemeinbevölkerung herangezogen hat.

Zum Schluss möchten wir noch vier Wirksamkeitsuntersuchungen analytischer Psychotherapie vorstellen, die dem Wirksamkeitsnachweis von vier analytischen Schulen dienen sollten. Es ging jeweils darum, die Effektivität der analytischen Behandlung durch in diesen Schulen ausgebildete Analytiker zu überprüfen.

### **7) Wirksamkeitsuntersuchung Adlerianischer Psychoanalysen (Rogner, 1994a u. b)**

Zur Untersuchung der Effekte von individualpsychologischer Psychotherapie hat Rogner 1994 eine Studie vorgelegt. Es wurden 45 Personen (Therapiegruppe-TG) ein Jahr nach Abschluss einer im Durchschnitt 237 Stunden umfassenden Adlerianischen analytischen Psychotherapie mit 38 Personen (Kontrollgruppe-KG), die eine Psychotherapie begonnen hatten, verglichen. Außerdem schätzten die 44 Personen der TG die Besserung ihrer bei Therapiebeginn angegebenen Beschwerden ein. Es wurde der „Gießen-Beschwerdebogen“, der „Gießen-Test“, das „Narzissmusinventar“ sowie der „Fragebogen zum emotionalen Befinden“ eingesetzt.

Ergebnis: Die TG litt weniger unter körperlichen Beschwerden, erlebte seltener negative und häufiger positive Emotionen und verfügte über eine effizientere narzisstische Regulation als die KG. Außerdem glich sie bevölkerungsrepräsentativen Stichproben (körperliche Beschwerden, psychosoziale Tendenzen) bzw. unterschied sich positiv von anderen klinischen Stichproben (emotionales Befinden, narzisstische Regulation).

Die Personen der TG hatten bei Therapiebeginn nach ihrer Einschätzung im Mittel unter 13 Beschwerden gelitten. Von diesen waren ein Jahr nach Therapieende neun stark gebessert. Die Besserungsrate bei den körperlichen Beschwerden war deutlicher (81 %) als bei den psychischen (65 %). Eine ausgeprägtere symptomatische Besserung war mit einer geringeren Belastetheit durch negative Emotionen und körperliche Beschwerden, einer größeren Stabilität des Selbstsystems und einer größeren Kontaktsicherheit und Offenheit in zwischenmenschlichen Beziehungen sowie einem günstigeren Therapieerfolg nach Einschätzung der Therapeuten assoziiert.

Methodische Mängel, insbes. die Einschätzung der Beschwerdeveränderungen über den langen Zeitraum von Behandlungsbeginn bis ein Jahr nach der Behandlung, statt zweier unabhängig voneinander vorgenommenen Einschätzungen, schmälern die Aussagekraft der Ergebnisse.

### **8) Die Wirksamkeit Jungianischer Psychoanalysen und Psychotherapien (Keller et al., 1997)**

Es handelt sich hierbei um eine katamnestiche Studie der DGAP. Das katamnestiche Design wird vor allem mit den langen Laufzeiten und den methodischen Schwierigkeiten anderer Designs begründet.

Auf der Basis freiwillig mitwirkender Therapeuten und Patienten wurden sechs Jahre nach Therapieende 111 Fälle untersucht, deren Behandlung 1987/1988 abgeschlossen war. Es wurden fünf Effektivitätskriterien herangezogen: 1. Die vergleichende Selbsteinschätzung der mitwirkenden Patienten hinsichtlich ihres Krankheitserlebens vor und nach Ende der Therapie, 2. die rückblickende Einschätzung des Gesundheitszustandes bei Therapieende durch die früheren Therapeuten, 3. Vergleich der mit den standardisierten Messverfahren SCL 90-R, Gießen-Test und VEV gemessenen Werte mit den Normwerten der gesunden Eichpopulation, 4. die Prä-Post-Beurteilung des aktuellen Krankheitsstatus mittels klinischer Nachinterviews an einer Teilstichprobe von 33 Patienten durch unabhängige Untersucher und 5. der Vergleich der bei den Krankenkassen erhobenen Arbeitsunfähigkeitszeiten und Krankenhausaufenthalte in einem Zeitraum fünf Jahre vor und nach der psychotherapeutischen Behandlung.

Im Ergebnis zeigte sich in allen gewählten Effektivitätsmaßen die Wirksamkeit jungianischer Therapien. In den standardisierten Messverfahren wurden Ergebnisse auf dem 0,01 Signifikanzniveau gemessen.

Grundsätzliche methodologische Einwände gegen die katamnestische Herangehensweise und der Verzicht auf die quasi-experimentelle Erhebung von Prä-Post-Daten beeinträchtigen die Aussagekraft der sehr guten Ergebnisse.

**9) Die Wirksamkeit der daseinsanalytischen Therapie bei psychosomatischen Störungen (Weber, 1994)**

Es handelt sich hierbei, wie die Autoren hervorheben, um die „erste empirische Arbeit zur daseinsanalytischen Psychotherapie“. (S. 69)

Die Untersuchung ist mit klarem Bezug und als Replik auf die Grawe'schen Veröffentlichungen angelegt und orientiert sich pragmatisch an der Methodik der empirischen Sozialforschung.

Die Untersuchung sucht denn auch den statistischen Beleg der Wirksamkeit der Daseinsanalyse zu führen. Dabei bedient sie sich 21 Falldarstellungen, die im Rahmen von Therapieausbildungen am Daseinsanalytischen Institut für Psychotherapie und Psychosomatik in Zürich in den Jahren 1974-1991 verfasst wurden, ohne dass hierbei bereits auf eine spätere statistische Auswertung rekurriert wurde. Die Autoren versuchen, diese Arbeiten als Datenbasis für eine Übersetzung in die Logik der empirischen Sozialforschung zu benutzen. Entsprechend stringent und methodologisch ‚sauber‘ formulieren sie zunächst falsifizierbare Ho-Hypothesen, werten dann das in eine empiristische Logik übersetzte Datenmaterial statistisch aus und kommen so dazu, die aufgestellten Ho-Hypothesen auf einem meist hohen Signifikanzniveau zurückweisen zu können. Nach Überführung der 21 Falldarstellungen in eine quasi-empiristische Datenbank wurden die folgenden „Variablen“ isoliert: Alter der Patienten, Geschlechtsverteilung, Behandlungsgrund, Kostenträger, Schichtzugehörigkeit, Therapeutengeschlecht, Therapeutenalter, Therapeutengrundberuf, Therapiesetting, Therapiedauer, Diagnose, Schweregrad der Erkrankung bei Therapiebeginn und Schweregrad bei Therapieende. Was die meisten dieser Variablen anbelangt, so bleibt die Studie in einer beschreibenden Darstellung verhaftet.

Lediglich der Grad der Erkrankung bei Therapiebeginn und bei Therapieende wird im statistischen Sinne als Prä-Post-Vergleich ausgewertet, unter Anwendung von t-Tests geprüft und in eine Berechnung der globalen Effektstärke vorangetrieben. Die Studie verzichtet aber darauf, über die globale Effektstärke hinaus zu Prüfungen des Zusammenhangs zwischen den isolierten Variablen und der Effektstärke zu kommen, - was sicherlich mit dem zu kleinen „N“ zu tun hat.

Trotz dieser recht empirisch anmutenden Angehensweise sind die Autoren selbstkritisch genug, die schwache empirische Ausgangslage und die möglichen wissenschaftslogischen Einwände nicht zu verschweigen. Besonders gravierend wirke sich der Umstand aus, dass die Veränderung der Krankheitsbilder in den untersuchten Falldarstellungen nur unzureichend dokumentiert wurde und nachträgliche durch Rater vorgenommene Einschätzungen doch recht kritikanfällig sind.

Die Studie kommt dennoch zu der globalen Aussage, dass es sich bei der Daseinsanalyse um eine Therapieform handele, die „statistisch abgesicherte“ und „hochsignifikante“ (S. 69) Therapieeffekte nachweisen kann.



**10) Langzeitwirkungen von Psychoanalysen und Psychotherapien: Eine multiperspektivische, repräsentative Katamnese studie** (Leuzinger-Bohleber, 2001)

Hierbei handelt es sich um eine rein katamnestiche Studie, die in den Jahren 1998-1999 durchgeführt wurde und in der alle Patienten rückwirkend untersucht wurden, die zwischen 1990 und 1993 ihre Behandlung bei einem DPV-Analytiker beendet hatten.

Ziel der Studie ist es, die Wirksamkeit von Langzeittherapien nachzuweisen und Aussagen über die Bedingungen der Wirksamkeit zu kommen.

Inhaltlich geht die Studie so vor, dass zunächst die zur Einschätzung des Therapieerfolges notwendigen Daten bei den untersuchten Patienten sowohl mit standardisierten Fragebögen als auch mit qualitativen Katamneseinterviews erhoben werden. In einem zweiten Auswertungsschritt werden die Daten zur Veränderungseinschätzung, zur Behandlungszufriedenheit und zur aktuellen Symptombelastung aus Sicht des Patienten, des ehemaligen Therapeuten und eines externen Experten herangezogen und der Grad der Übereinstimmung zwischen diesen Einschätzungen ermittelt, um die Aussagen gegenseitig zu validieren. Die so ermittelten „Therapieerfolge“ werden durch eine vergleichende Heranziehung des objektiven Datums der Inanspruchnahme von allgemeinmedizinischen Gesundheitsleistungen noch einmal extern validiert.

In einem zweiten Studienteil wird durch eine Verzahnung von quantitativen Daten und qualitativen Katamneseinterviews eine taxonomische Clusterbildung im Sinne einer klinischen Typenbildung vorgenommen, welche die unterschiedlichen Muster des langfristigen Umgangs mit der Therapieerfahrung darstellen sollen. Auf diese Weise wurden acht unterschiedliche Typen ermittelt, beschrieben, inhaltlich definiert, prototypisiert und durch externe Rater validiert.

Andere Aussagen der Studie, dass der Erfolg der Therapien sehr stark vom Matching-Prozess zwischen Patient und Therapeut abhängt, dass die nachhaltige Wirkung der Therapien nicht unwesentlich von der patientenzentrierten Beendigung der Therapie abhängt, dass hochfrequente Therapien differenziertere Erfolge erzielen, dass mangelnder Therapieerfolg etwas mit der nicht aufgelösten negativen Übertragung zu tun hat, sowie Einsichten in die Bedeutung der postanalytischen Phase entstammen den Einschätzungen der Rater und Forscher, sind aber aus empiristischer Sicht in dieser Studie nicht objektiv überprüfbar belegt.

Der gesamtmethodische Ansatz dieser Studie wird von den Autoren wie folgt begründet: „Wie wir an anderer Stelle diskutiert haben, entsprechen wir mit diesem Vorgehen dem Stand aktueller wissenschaftstheoretischer Überlegungen, die von einem einheitswissenschaftlichen Verständnis von ‚wissenschaftlicher Methodik und Wahrheitsüberprüfung‘ im Sinne des logischen Empirismus weit abgerückt sind und die Spezifität verschiedener Forschungsmethoden und Prüfkriterien in den verschiedensten wissenschaftlichen Disziplinen ins Zentrum der kritischen Reflexionen stellen“. (Leuzinger-Bohleber et al., 2001, S.196)

Die Studie legt sehr viel Wert auf die Sicherstellung der Repräsentativität der untersuchten Patienten und zieht zur Sicherstellung der externen Validität die bei den Krankenkassen erfragbaren objektiven Daten der Inanspruchnahme von Gesundheitsleistungen heran.

Wesentliche Kriterien empirischer Methodologie sind nicht berücksichtigt worden, wie z.B. der Prä-/Post-Vergleich, die quasi-experimentelle Kontrolle mittels einer Kontrollgruppe, das Fehlen einer Randomisierung u.a.

Von Rad et al. (2001, S. 314) formulieren ihre Stellungnahme so: „Das DPV-Katamnese-Schiff von Leuzinger-Bohleber hat die Scylla der untersuchungsbedingten Verfälschung der untersuchten Therapieverfahren ernst genommen und weiträumig umfahren. Damit hat es sich aber dem fast unwiderstehlichen methodischen Strudel der Charybdis internationaler Forschungs-Standards gefährlich genähert, der es zu verschlingen droht. Vielleicht will es dem ja entschlossen trotzen, wie auch die ganze Anlage der Studie entschlossen die psychoanalytische Fahne gehisst hat und sich indirekt wohl mehr an die Psychoanalytiker als an die „scientific community“ der Therapie-Forscher wendet.“

Abschließend ist festzustellen, dass die referierten neueren Forschungsbemühungen der Phase drei wenig zur Beantwortung der Frage, was wirklich im Therapieprozess geschieht und für die Wirksamkeit der Behandlungen ausschlaggebend ist, beigetragen haben. Die Untersuchungen, insbesondere nach der Grawe-Veröffentlichung 1994, kreisen weiterhin um die Frage der Wirksamkeit psychoanalytisch begründeter Verfahren, und wie diese mit über die Grenzen der psychoanalytischen Gemeinschaft hinaus anerkannter sozialwissenschaftlicher Methodik nachgewiesen werden kann. Der Legitimationsdruck angesichts knapper gewordener finanzieller Mittel, der Zwang, die Effektivität des eigenen Verfahrens auf wissenschaftlich anerkannte Weise darzulegen, bestimmt den Stand der gegenwärtigen Psychotherapieforschung. Insofern stehen Diskussionen um die „richtige“ Methode in der wissenschaftlichen Auseinandersetzung sehr viel mehr im Vordergrund als Anstrengungen, die mikroanalytischen, prozessualen Bedingungen für erfolgreiche und weniger erfolgreiche Behandlungen genauer zu untersuchen und damit neue, in der Praxis des Psychoanalytikers umsetzbare Erkenntnisse zu gewinnen.

### 3. Wissenschaftstheorie und Psychoanalyse

In der psychoanalytischen Literatur und mit ihr auch in der Literatur zur Individualpsychologie fällt der Mangel an wissenschaftstheoretischen Darlegungen auf. Das soll nicht heißen, dass die Psychoanalyse keine implizite Wissenschaftstheorie hat, sondern soll lediglich heißen, dass diese Wissenschaftstheorie mit dem ‚fallbezogenen Verfahren‘ so verschmolzen ist, dass sie nicht als eigenständiges System unabhängig von der fallbezogenen Inhaltlichkeit psychoanalytischer Aussagen erkennbar ist.

Wissenschaftstheoretische Meta-Überlegungen scheinen eine Angelegenheit des universitären Lebens zu sein, wohl weil man sich dort immer schon denn fall-losgelösten Luxus leisten konnte, sich auf eigens dazu eingerichteten Lehrstühlen Gedanken darüber machen zu können, welche Kriterien denn erfüllt sein müssen, um das eigene Tun als „wissenschaftlich“ bezeichnen zu können. Bis vor einigen Jahren fehlte im Bereich der außeruniversitären Psychoanalyse jedweder ernsthafte Versuch, sich an der im universitären Bereich geführten wissenschaftstheoretischen Diskussion zu beteiligen. Auf die weitgehende Spaltung von Psychoanalyse und universitärem Wissenschaftsbetrieb und die sich daraus ergebende wissenschaftstheoretische Abschottung der Psychoanalyse weist besonders Kaiser (1993, S. 354) hin. Mertens fragt mehr aus einem Unverständnis heraus, warum die Psychoanalyse als „eine weltweit angesehene Disziplin von den deutschen psychologischen Fakultäten so gut wie nicht zur Kenntnis genommen wird“ (Mertens, 1994, S. 90). Eine Erklärung sucht Mertens zum einen in politisch-geschichtlichen Gründen wenn er vermutet, dass dies „speziell mit den Deutschen“ (Mertens, 1994, S. 90) und ihrer unbewältigten Vergangenheit zu tun habe, zum anderen nimmt er Zuflucht zu der Vermutung, dass der Psychoanalyse der Zugang zu den Universitäten versperrt sei, weil die „ausschließliche(n) Universitätsforscher ... in der Regel den Bezug zur Praxis“ (Mertens, 1994, S. 90) verlor und die Leistung der Psychoanalyse deshalb nicht mehr wahrnehmen könnten. Solche Überlegungen müssen inhaltlich ja nicht falsch sein, lenken aber von den tatsächlichen Versäumnissen der Psychoanalyse auf wissenschaftstheoretischem Gebiet ab, - Versäumnisse, die der Psychoanalyse vielleicht ihre universitär-wissenschaftliche Relevanz gekostet haben. In diesem Zusammenhang ist es z.B. auffällig, dass die Psychoanalyse die hauptsächlich im Bereich der universitären Soziologie und Philosophie unter dem Begriff „Positivismusstreit“ von Popper, Albert, Habermas und Adorno (vgl. Th. W. Adorno u.a., 1972: „Der Positivismusstreit in der deutschen Soziologie“) geführte wissenschaftstheoretische Auseinandersetzung der letzten Jahrzehnte unbeteiligt hat an sich vorbeigehen lassen, obwohl deutliche Angebote an die Psychoanalyse ergangen sind, sich an der Diskussion zu beteiligen. Beispielhaft sei hier auf einen der bedeutendsten Wortführer des „Positivismusstreits“ und Vertreter der „Frankfurter Schule“, den Soziologen Jürgen Habermas verwiesen, der in seinem 1982 erschienenen Buch „Zur Logik der Sozialwissenschaften“ die Psychoanalyse sogar gegen den methodischen Empirismus Poppers ins Feld führt, wenn er schreibt:

„Die Erfahrungsbasis der strikten Wissenschaften ist nicht unabhängig von den Standards, die diese Wissenschaften selber der Erfahrung anlegen. Offensichtlich ist die Prozedur der Überprüfung, die Albert als die einzig legitime unterstellt, nur eine unter mehreren. Moralische Gefühle, Entbehrungen und Frustrationen, lebensgeschichtliche Krisen, Einstellungsänderungen im Zuge einer Reflektion – vermit-

teln andere Erfahrungen. Sie können durch entsprechende Standards zur Prüfungsinstanz erhoben werden; die vom Psychoanalytiker ausgewertete Übertragungssituation zwischen Arzt und Patient gibt dafür ein Beispiel“. (Habermas, 1982, S. 48)

Allerdings fällt Habermas hinter sich selbst zurück, wenn er im empirisch verhafteten Sinne empfiehlt,

„die psychoanalytischen Annahmen im kategorialen Rahmen einer strikten Erfahrungswissenschaft zu reformulieren. So sind Theoreme im Rahmen der behavioristisch angelegten Lernpsychologie neu formuliert und dann den üblichen Prozeduren der Überprüfung unterzogen worden. Anspruchsvoller ist der Versuch, das ichpsychologisch fortgebildete, aber triebdynamisch begründete Persönlichkeitsmodell mit Mitteln des neueren Funktionalismus als ein selbstgeregeltes System zu rekonstruieren. In beiden Fällen ermöglicht der neue theoretische Rahmen eine Operationalisierung der Begriffe, in beiden Fällen verlangt er eine Überprüfung der abgeleiteten Hypothesen unter experimentellen Bedingungen.“ (Habermas, 1973, S. 308 f.)

Man muss aber auch sehen, dass sich trotz solcher Rückfälle bei Habermas erste Ansätze zur Begründung einer eigenständigen psychoanalytischen Wissenschaftstheorie finden. In seinem Buch „Erkenntnis und Interesse“ widmet Habermas der Psychoanalyse als Wissenschaft ein komplettes Kapitel mit der Bezeichnung „Selbstreflexion als Wissenschaft: Freuds psychoanalytische Sinnkritik“ und spricht das wissenschaftstheoretisch Neue der Psychoanalyse als etwas an, das zwar etwas mit Hermeneutik und auch etwas mit Naturwissenschaften zu tun hat, das aber dennoch etwas anderes, man könnte sagen etwas ‚Neues‘, und zwar ‚etwas bisher nicht umgesetztes Neues‘ ist:

„Mit der Entstehung der Psychoanalyse eröffnet sich die Möglichkeit eines methodologischen, von der Logik der Forschung selbst gebahnten Zugangs zu jener vom Positivismus verschütteten Dimension. Diese Möglichkeit ist nicht realisiert worden, denn das szientistische Selbstmissverständnis der Psychoanalyse hat jene Möglichkeit verstellt. Freilich ist das Missverständnis nicht ganz unbegründet. Die Psychoanalyse verbindet nämlich Hermeneutik mit Leistungen, die genuin den Naturwissenschaften vorbehalten zu sein scheinen.“ (Habermas, 1973, S. 262 f.)

Und weiter:

„Allgemeine Interpretationen teilen nämlich mit allgemeinen Theorien den weitergehenden Anspruch, kausale Erklärungen und bedingte Prognosen zu gestatten. Aber im Unterschied zu den strikten Erfahrungswissenschaften kann die Psychoanalyse diesen Anspruch nicht auf der Grundlage einer methodologisch klaren Trennung des Objektbereichs von der Ebene theoretischer Aussagen einlösen. Das

hat Konsequenzen für den Aufbau der Interpretationssprache, für die Bedingungen der empirischen Überprüfung und schließlich für die Logik der Erklärung selber.“ (Habermas, 1973, S. 320)

Die Habermas'schen Überlegungen gehen eindeutig in wissenschaftstheoretische Richtung, wenn er von der besonderen „methodologischen Eigentümlichkeit“ (Habermas, 1973, S. 324) der Psychoanalyse spricht, nämlich

„dass allgemeine Interpretationen nicht den gleichen Kriterien der Widerlegung gehorchen wie allgemeine Theorien. Wenn eine bedingte Prognose, die aus Gesetzhypothese und Ausgangsbedingungen abgeleitet ist, falsifiziert wird, darf die Hypothese als widerlegt gelten. Eine allgemeine Interpretation können wir analog überprüfen, indem wir aus einer ihrer Ableitungen und den Mitteilungen eines Kranken eine Konstruktion ableiten. Dieser Konstruktion können wir die Form einer bedingten Prognose geben. Wenn sie zutrifft, wird der Patient dazu gebracht, bestimmte Erinnerungen zu reproduzieren, .... Aber der Weg der Falsifikation ist in diesem Falle nicht derselbe wie bei allgemeinen Theorien. Denn wenn der Kranke eine Konstruktion ablehnt, gilt die Interpretation, aus der sie abgeleitet worden ist, noch keineswegs als widerlegt. Psychoanalytische Annahmen beziehen sich nämlich auf Bedingungen der Suspendierung eben der Erfahrung, an der sie sich doch bewähren müssen: die Erfahrung der Reflexion ist die einzige Instanz, an der Hypothesen sich bewähren oder scheitern können. Wenn sie ausbleibt, besteht immer noch die Alternative: entweder ist die Interpretation falsch oder aber die im übrigen richtig diagnostizierten Widerstände sind zu stark.“ (Habermas, 1973, S. 324 f.)

Uns sind keine eigenen Beiträge von psychoanalytischer Seite zu der von Habermas angesprochenen wissenschaftstheoretischen Relevanz der psychoanalytischen Gegenstandsbildung bekannt. Scheinbar findet sich kein Vertreter der Psychoanalyse, der die diesbezüglichen Habermas'schen Gedanken aufzugreifen, weiterzuführen oder überhaupt der Erwähnung wert gefunden hätte. So wurde die Diskussion um die wissenschaftstheoretischen Potentiale der Psychoanalyse an den Universitäten lange Zeit ohne Beteiligung der Psychoanalyse geführt, - eine für die Psychoanalyse verpasste Chance, sich wissenschaftstheoretisch zu profilieren und sich jenseits von Hermeneutik und Empirismus wissenschaftstheoretisch eigenständig zu positionieren.

Erst seit Grawe ist ein intensiveres Bemühen der Psychoanalyse um wissenschaftstheoretische Positionen zu bemerken. Im Bereich der Individualpsychologie sind diesbezüglich die Autoren Heckrath und Dohmen (1994), Brunner (1994) sowie Presslich-Titscher und Datler (1994) zu nennen. Im weiteren psychoanalytischen Bereich ist hier besonders auf die oben mit ihren Publikationen bereits genannten Autoren Kaiser und Mertens zu verweisen. Alle diese Autoren mischen sich zwar in die wissenschaftstheoretische Diskussion mit Grawe ein, bleiben jedoch letztlich den bereits bekannten Argumenten der Auseinandersetzung von Hermeneutik und naturwissenschaftlich orientierten Empirismus verhaftet. Die wissenschaftstheoretische Wortmeldung von Kaiser macht auf den „normativen Schritt“ aufmerksam, welcher der unhinterfragten Anwendung

des Empirismus auf die Psychoanalyse zugrunde liegt. Die Geltung des „hypothetisch-deduktiven Induktivismus“ und des „logisch-empirischen Wissenschaftsverständnisses“ für die Psychoanalyse werde von Grawe wie selbstverständlich unterstellt, wohingegen diese Geltung doch eigentlich „seinerseits erst zu legitimieren wäre“ (Kaiser, 1993, S. 348).

Einen anderen und wirklich weiterführenden Ansatz liefern dagegen Leikert, Ruff und Berleburg (1994), die, - und das ist (unter Außerachtlassung der referierten Habermas'schen Überlegungen) wirklich neu -, auf die Möglichkeit und Notwendigkeit einer völlig eigenen psychoanalytischen Wissenschaftstheorie aufmerksam geworden sind:

„Zur Charakterisierung der psychoanalytischen Methodik reicht nicht aus zu sagen, sie sei beschreibend; denn sie beschreibt das, was sie erfasst in einer bestimmten Perspektive, nämlich auf den unbewussten Konflikt hin.

Ausgehend von einem fast beliebigen szenischen Element, welches der Patient anbietet, werden Hypothesen bezüglich der zugrundeliegenden unbewussten Szene gebildet. Diese werden dann am weiteren Material geprüft, modifiziert und so lange weiterverfolgt, bis eine ausreichende Erkenntnissicherheit erreicht ist.

Betrachtet man dieses Vorgehen auf einer formalen Ebene, so werden darin die systematische Wiederholung und die Variation der Beobachtung deutlich, die also ... übergeordnete Prinzipien für wissenschaftliche Hypothesenprüfung sind.

Die erste Phase der Untersuchung ist abgeschlossen, wenn eine genaue und damit prüfbare Beschreibung eines Typus vorliegt. Jetzt wird diese Hypothese an weiteren Einzelfällen hinsichtlich ihrer Angemessenheit an die je individuelle Entwicklung und ihres Geltungsbereichs, also hinsichtlich ihrer Verallgemeinerungsfähigkeit, geprüft.

Dieses Verfahren ist keineswegs originell; denn in gewisser Weise ist die psychoanalytische Erkenntnis immer auf diese Weise vom Besonderen zum Allgemeinen fortgeschritten.

Um so erstaunlicher ist es, dass bisher nie versucht wurde, diese beschreibende Verfahren in einem methodischen, genau definierten Rahmen einzusetzen; denn nur ein solcher Rahmen ermöglicht es, die Ungenauigkeiten der Intuition zu vermeiden. Nur ein systematischer Vergleich mit genau definierten Untersuchungsgruppen erlaubt es, zu Aussagen zu kommen, die wissenschaftlichen Ansprüchen gerecht werden.“ (Leikert u.a., S. 79 f.) (Unterstreichungen nicht im Original.)

Was Leikert u.a. hier bemerken ist, dass die Psychoanalyse zwar immer schon nach bestimmten Ableitungsprinzipien (Methodologie) gearbeitet hat, dass diese Ableitungsprinzipien jedoch nicht ausdrücklich ausgesprochen, aufgeschrieben, zu einem begrifflichen System verbunden und zu einer wissenschaftstheoretischen Konstruktion weiterentwickelt wurden. Es reicht nämlich im heutigen Wissenschaftsbetrieb nicht mehr aus zu sagen, man habe eine Methode, und es reicht auch nicht aus, das eigene Forschungsverfahren von einer impliziten (sozusagen wissenschaftlich unbewussten) Wissenschaftstheorie organisieren zu lassen

(wie z.B. in den Falldarstellungen der Psychoanalyse), sondern man muss diese Wissenschaftstheorie und die aus ihr abgeleiteten methodischen Schritte (Ableitungsregeln der Aussagen) explizit benennen und beschreiben können. Und es reicht auch nicht aus, wenn einzelne methodische Aspekte in psychoanalytischen Veröffentlichungen gelegentlich einmal erwähnt werden, sondern um eine Wissenschaftstheorie zu begründen ist es nötig, ein komplettes System, also den organisierten und organisierenden Zusammenhang von Ableitungsregeln benennen und aufschreiben zu können. Auch der allgemeine Hinweis auf die wissenschaftstheoretische Verankerung der Psychoanalyse in der Hermeneutik hilft hier wenig, weil die Übersetzung in die psychoanalytische Gegenstandsbildung und die Umsetzung in wiederholbare methodische Schritte fehlt. Weder die Hermeneutik noch die Soziologie, - und das mag man an den doch eher etwas holprig wirkenden Versuchen von Habermas erkennen -, sind in der Lage, diese Aufgabe der Übersetzung und Ausformulierung zu einer psychoanalytischen Wissenschaftstheorie zu leisten.

Dies muss die Psychoanalyse schon selbst hervorbringen, wobei die Überlegungen von Leikert u.a. hierzu einen ersten Zugang darstellen können, die sich im übrigen sehr treffend in die Habermas'schen Überlegungen einfügen.

Solange es der Psychoanalyse jedoch noch nicht gelungen ist, eine eigene wissenschaftstheoretische Position der Psychoanalyse schriftlich umfänglich und zergliedert zu begründen, solange ist sie in gewissem Sinne in der schwächeren Position und gibt der empirischen Wissenschaftstheorie aus deren Sicht die Legitimation, auf dem – wie Habermas es nennt - „Programm der Einheitswissenschaft“ (Habermas, 1972, S. 157) zu bestehen. Diese kann nämlich nicht nur auf Empirie und Falsifikationstheorem, sondern sie kann auch auf ein aufgeschriebenes System ihrer Ableitungsregeln zur Gewinnung und Begründung ihrer Aussagen verweisen, deren Nichteinhaltung Begründung dafür ist, Aussagen als – aus ihrer Sicht – methodisch nicht begründet und damit als unwissenschaftlich abzuweisen. Entweder ist die Psychoanalyse langfristig in der Lage, eine eigene ausformulierte Wissenschaftstheorie der Psychoanalyse zu entwickeln, oder aber sie beugt sich dem empirischen Positivismus und liefert im empiristischen Sinne methodologisch ‚saubere‘ Wirksamkeitsuntersuchungen ab. Zur Zeit fehlt beides, sowohl die konsequent eigenständige Wissenschaftstheorie und der aus ihr abgeleitete Kanon von Ableitungsregeln wissenschaftlicher Wirksamkeitsuntersuchungen als auch der konsequente Versuch eines empirisch-positivistischen Wirksamkeitsnachweises.

Im Rahmen dieser Problemanzeige möchte die nachfolgende Wirksamkeitsuntersuchung ihre wissenschaftsgeschichtliche Platzanweisung finden. Zum einen geht es in Hinblick auf Grawes schonungslos vernichtende Bewertung der „Individualpsychologisch-Psychoanalytischen Psychotherapie“ um eine exklusive Untersuchung ausschließlich individualpsychologischer Psychotherapien, und zum anderen geht es darum, einen Wirksamkeitsnachweis für die „Individualpsychologisch-Psychoanalytische Psychotherapie“ unter konsequenter Beachtung aller empirisch-positivistischen Methodologiekriterien vorzulegen. Die Einhaltung des Falsifikationstheorems sollte dabei an vorderer Stelle wichtig sein, um dem „Gegenstand“ die Möglichkeit einzuräumen, die Wirksamkeit der „Individualpsychologisch-Psychoanalytischen Psychotherapie“ widerlegen zu können. Dagegen steht allerdings unsere nicht verleugnete Vermutung, dass die „Individualpsychologisch-

Psychoanalytische Psychotherapie“ wirksam ist und dass diese Wirksamkeit auch mit Verfahren des methodischen Empirismus nachweisbar ist, obwohl diese Verfahren dem psychischen Gegenstand aus psychoanalytischer Sicht nicht ganz angemessen sind. Diese Position gilt es zu begründen: Die Nichtangemessenheit besteht darin, dass es beim psychoanalytischen Geschehen um Veränderungen auf der Strukturebene geht, während die empiristische Methodologie doch lediglich auf der Symptomebene anzusetzen in der Lage ist. Die Strukturebene aber ist „mehr und anders als die Summe der Teile“, - hier mehr und anders als die Summe der Symptome. Wenn es um die Strukturebene der Erfassung von psychisch-dynamischen Gesamtzusammenhängen geht, dann ist der Empirismus dem psychoanalytischen Modell weit unterlegen. Die eingengte messtheoretische Erfassung von empirisch Erfahrbarem reicht an die Erkenntnisgewinnung durch einen „Prozess des szenischen Verstehens und der Teilhabe an den unbewussten Inszenierungen des zu beforschenden Gegenübers“ (Mertens, 1994, S. 81) unter Einbeziehung der Gegenübertragung bei weitem nicht heran. Aus psychoanalytischer Sicht bleiben die empiristischen Verfahren der Symptomebene verhaftet, ohne auf die Strukturebene und die sich dort abspielenden qualitativen Veränderungen durchgreifen zu können. Langenmayr und Kosfelder (1998) sprechen in diesem Zusammenhang von der Notwendigkeit, die „Ebenen der Evaluation von Psychotherapie“ (S. 97) zu unterscheiden, - meint hier die Unterscheidung der Ebene der Symptomatik von der Ebene der Struktur bzw. der Krankheitsursachen. Das Starren auf die „Veränderungen auf der Symptomebene“ ohne Berücksichtigung der sie bedingenden „Veränderungen auf der Strukturebene“ erscheint dem psychoanalytischen Denken als eine höchst laienhafte und naive Angehensweise. Die qualitativen Veränderungen, die u.U. zwar keine quantitative Ausprägung haben, aber je nach „Fall“ den größten Teil der Veränderungen darstellen, bleiben von vorne herein unberücksichtigt. Und was die quantitativen Ausprägungen qualitativer Veränderungen anbelangt, so bleibt deren Einbettung und Bedeutung im Rahmen der strukturellen Gegebenheiten völlig unerforscht. Derart von vorne herein sowohl im Forschungsgegenstand wie auch im zu erwartenden Erkenntnisgewinn eingengt, verfolgt die nachfolgende Untersuchung den Versuch, auf der Symptomebene eine quantitative Spurensuche für viel umgreifendere qualitative Veränderungen zu betreiben, welche auf der Strukturebene liegen, von dieser Studie jedoch nicht erfasst werden können.

#### **4. Empiristisch-wissenschaftslogische Kriterien**

Wir haben für die nachfolgende Untersuchung von Anfang die Position eingenommen, die Wirkung analytischer Therapien nach den methodischen Regeln des Empirismus untersuchen zu wollen. Es reizte uns die Frage, welche Resultate sich zeigen würden, wenn wir die psychoanalytischen Therapien, und hier insbesondere „Individualpsychologisch-Psychoanalytische Psychotherapien“, einmal mit den auf diesen Gegenstand nie wirklich konsequent angewandten Regeln der empirischen Sozialforschung untersuchen würden. Wir waren uns darüber klar, dass wir uns damit aus den psychoanalytischen Reihen dem Vorwurf aussetzen würden, „die Psychoanalyse zu Grawe zu tragen“. (Heinzel, 1996, S. 7)



Wir entschieden uns, die Kriterien der Zählbarkeit, der Messbarkeit und das Studiendesign der „Methodik der Empirischen Sozialforschung“, als deren Vertreter in Deutschland z.B. Renate Mayntz u.a. („Einführung in die Methoden der empirischen Soziologie“, 3. Aufl., Opladen 1972) genannt werden kann, einmal möglichst konsequent auf „Individualpsychologisch-Psychoanalytische Psychotherapien“ anzuwenden. Und wir nahmen uns vor, von vorne herein einmal alles zu unterlassen, was uns im Nachhinein aus dem Lager der Empiristen als methodologische Einwendung hätte vorgehalten werden können. Dies hatte folgende Konsequenzen für unser Studiendesign:

- a. Bildung von Untersuchungs- und Kontrollgruppe,
- b. Zufallsauswahl unserer Populationen,
- c. Einsatz von im empiristischen Sinne objektiven und vom Forscher unabhängigen Messverfahren,
- d. vorrangige Verwendung von ‚hier-und-jetzt‘ bezogenen objektiven Befindlichkeitsskalen,
- e. Handhabung von Rohwerten als objektive „Daten“,
- f. Verzicht auf alle Fremdeinschätzungen durch Therapeuten und Untersucher, sondern Datenerhebung ausschließlich bei den Patienten,
- g. Trennung von Therapeut und Forscher, d.h. keine Teilnahme von Patienten der Forscher (die gleichzeitig Psychoanalytiker sind) an der Untersuchung,
- h. Raum-zeitliche Trennung von therapeutischer Situation und Datenerhebung, um die Auswirkungen von „Übertragungskonstellationen“ so gering wie möglich zu halten,
- i. Anonymität der Patienten, um diesen die Möglichkeit zu sichern, sich sowohl hinsichtlich der Therapeuten wie auch hinsichtlich der Forscher völlig frei von Erwartungsvermutungen äußern zu können,
- j. Veränderungsmessungen durch den Vergleich wiederholter Messungen zu fünf verschiedenen Messzeitpunkten (Prä-Post-Messungen zu fünf verschiedenen Messzeitpunkten t1 bis t5) mit den gleichen Messinstrumenten.

Die Einhaltung dieser methodischen Grundsätze schien uns nötig, um eventuellen Vorwürfen vorzubeugen, die durchführenden Untersucher hätten durch methodische Manipulationen Einfluss auf die Ergebnisse der Studie im Sinne der Erzeugung gewünschter Resultate genommen.

Heinzel, Breyer und Klein, selbst psychoanalytisch orientiert, haben in ihrer Veröffentlichung (1996) „Ambulante Psychoanalyse in Deutschland“, eine wohl relativ vollständige Liste methodischer Einwände/Kriterien zusammengestellt, die zu beachten jeder psychoanalytischen Wirksamkeitsforschung nahegelegt werden muss, um in der Frage der Prüfung ihrer Wissenschaftlichkeit vor empiristischer Kritik bestehen zu können. Mit gleicher Zielsetzung, aber etwas systematischer, haben Langenmayr und Kosfelder (1998) die Kriterien der „Evaluation in der Psychotherapie“ zusammengestellt. Beide Veröffentlichungen beschreiben in wesentlichen Grundzügen gleiche Rahmenbedingungen der Wissenschaftlichkeit von Psychotherapieforschung. Im Einzelnen handelt es sich hierbei um die folgenden Punkte:

a) „Der Therapeut weiß in der Regel (und der Patient bisweilen ebenfalls), dass er an einer Evaluationsstudie teilnimmt, wodurch Ablauf und Ergebnis der Therapie beeinflusst werden können“ (Heinzel u.a., 1996, S. 2). Langenmayr und Kosfelder (1998) sprechen in diesem Zusammenhang vom „sog. Hawthorne-Effekt“, der besagt, „dass das Wissen um die Teilnahme an einer Evaluationsstudie zu erfolgsrelevanten Veränderungen führt“ (S.105). Dieser Einwand trifft natürlich auch auf unsere Untersuchung zu, - allerdings, - und das ist nicht ganz unwichtig -, sowohl für die Untersuchungs- als auch für die Kontrollgruppe. Dadurch, dass es sich um einen auf beide Gruppen auswirkenden Einflussfaktor handelt, wird aus ihm ein systematischer Messfehler, der aber wegen des Systematischen seines Einflusses in unserer Untersuchungsgruppen-Kontrollgruppen-Vergleichsmessung vernachlässigt werden kann, weil ja der ‚Abstand‘ der Mittelwerte gleich bleibt, wenn beide Mittelwerte gemeinsam entweder nach oben oder nach unten verschoben werden. So kann als gesichert angenommen werden, dass ein eventuell gemessener Wirkungsunterschied zwischen Untersuchungs- und Kontrollgruppe nicht mehr auf die Variable „Evaluationsstudie“, sondern, - ceteris paribus, die Gleichheit bzw. Kontrolle aller anderen Wirkfaktoren unterstellt -, nur noch auf die intervenierende Variable „Psychotherapie“ bzw. „keine Psychotherapie“ zurückführbar ist. Unabhängig von diesem Argument wäre es natürlich auch ein spektakuläres Ergebnis, wenn nachweisbar wäre, dass klinisch relevant erkrankte Patienten allein schon durch die Teilnahme an einer Evaluationsstudie dauerhaft zu heilen sind, - hieraus ließen sich zwar merkwürdig anmutende, gleichwohl aber interessante Therapiekonzepte ableiten.

b) „Die Erfolgsindikatoren werden nicht von den behandelnden Medizinern definiert, sondern von den neutralen Forschern vorgegeben“ (Heinzel u.a., 1996, S. 2). Dieses Kriterium ist in unserer Untersuchung insofern erfüllt, als dass die Therapeuten an der Auswahl der Messinstrumente, der methodischen Bearbeitungsschritte und der Ergebnisinterpretation nicht beteiligt waren. Es kann jedoch der Gegeneinwand kommen, dass die „Forscher“ selbst „behandelnde Therapeuten“ sind, und insofern die Auswahl der Messinstrumente nicht unvoreingenommen stattfand. Hierzu ist zu sagen, dass keine „selbst entwickelten Erfolgsdefinitionen“ verwendet wurden, sondern dass unter Zuhilfenahme anerkannter Testverfahren wie den Gießener Beschwerdebogen (GGB), den Minnesota Multiphasic Personality Inventory, deutsche Kurzform (im Folgenden abgekürzt als MMPI/K) und dem Veränderungsfragebogen des Erlebens und Verhaltens (VEV) sowohl objektive Prä-Post-Vergleiche und als auch vergleichende Gegenüberstellungen mit externen Validierungswerten durchgeführt wurden, um die Ergebnisse an geltenden Standards zu messen, die unabhängig von dieser Studie entwickelt und anerkannt sind. Langenmayr und Kosfelder (1998) plädieren dafür, dass „unter den Messinstrumenten ... sich immer solche befinden (sollten), über die ein allgemeiner Konsens besteht“ (S. 107). Indem wir uns an dieses Kriterium gehalten haben, glauben wir den Verdacht der subjektiven Einflussnahme der Forscher entkräften zu können. Die Entscheidung für den VEV beruhte im übrigen darauf, dass dieses Verfahren von Grawe selbst, also dem großen Zweifler an der „Individualtherapie“, als Maßstab für die Messung von Psychotherapieerfolgen genannt wird.

Auch wurde auf Datenverarbeitungsprogramme (Word und SPSS) und statistische Signifikanztestverfahren (t-Tests und Varianzanalyse) zurückgegriffen, die ebenfalls nicht von den „Forschern“ und gemeinhin vom positivistischen Wissenschaftsbetrieb anerkannt sind. Die eigentliche statistische Auswertung mit Hilfe des

SPSS wurde von einem unabhängigen Statistiker und Nicht-Psychotherapeuten auf Honorarbasis vorgenommen. Bleibt der Hinweis, dass die Forscher natürlich keine eigenen Patienten in die Untersuchung eingebracht haben, um eventuellen methodologischen Zweifeln auch von dieser Seite zu begegnen.

c) „Die Forschung orientiert sich am subjektiven Befinden der Patienten“ (Heinzel u.a., 1996, S.2). So einfach diese Formulierung erscheint, in ihr stecken zwei Erfordernisse: Zum einen nämlich, dass nur ‚Patienten‘ und nicht etwa die Therapeuten befragt werden, und zum anderen, dass primär möglichst das tatsächliche jeweilige positiv vorliegende subjektive ‚Befinden‘ und nur sekundär die retrospektive Einschätzung des Therapieerfolgs erfragt wird. Beide Erfordernisse haben wir zu erfüllen gesucht:

So haben wir nicht die Therapeuten befragt, z.B. etwa nach ihrer Einschätzung hinsichtlich des Befindens der Patienten oder ihre Einschätzung hinsichtlich des vermeintlichen Therapieerfolgs, sondern wir haben ausschließlich die Patienten befragt, denn nur diese können sinnvoller Weise Auskunft über ihren Zustand geben. Dem zweiten Erfordernis haben wir dadurch versucht gerecht zu werden, dass wir der Frage nach den tatsächlich positiv vorhandenen Krankheitssymptomen zum jeweiligen Messzeitpunkt den Vorrang vor der Frage nach einer Therapieerfolgseinschätzung gegeben haben. Langenmayr und Kosfelder (1998) machen darauf aufmerksam, dass „mit einem Auseinanderklaffen von ... Selbstbeurteilungen der KlientInnen und objektiven Therapieerfolgsmaßen gerechnet werden“ (S. 100) muss und die „retrospektive Einschätzung ... Verzerrungen (unterliegt)“ (S.103), weshalb wir am liebsten auf die Verwendung solcher Einschätzskalen völlig verzichtet hätten. Dass wir uns dann doch für die Verwendung des VEV entschieden haben, hat damit zu tun, dass Grawe dieses Verfahren in seinem Forschungsgutachten besonders wertschätzt und uns bei Durchsicht der Literatur aufgefallen war, dass hinsichtlich dieses Verfahrens ein allgemein „wertschätzender Konsens“ vorherrscht. Dass wir im katamnestischen Teil unserer Studie auf den Helping Alliance Questionnaire (HAQ) und die Intrex-Kurzform, - beide beinhalten retrospektive Erfolgseinschätzungen -, zurückgegriffen haben, resultiert aus erweiterten Fragestellungen, die mit tiefenpsychologischen Konzepten (vgl. hierzu Kap. 11.2.3.1.-2.) zu tun haben. Der eigentliche Kern unserer Studie wird jedoch mit Hilfe des MMPI(K) und des GBB bestritten. Beide Verfahren verlangen vom Probanden keine retrospektive Einschätzung, sondern aktuelle Befindlichkeitsskalierungen zum jeweilig abgefragten Messzeitpunkt. Den Vergleich zu vorherigen Messzeitpunkten führt nicht der Patient durch, sondern dies machen wir, indem wir seine Aussagen zu den verschiedenen Messzeitpunkten miteinander vergleichen. So vermieden wir es, dass der Patient in die Not kommt, sich erinnern zu müssen, was er zu vorherigen Zeitpunkten auf die einzelnen Items geantwortet hat.

Wir versuchten so der Tendenz entgegenzuwirken, dass die Patienten in dem Versuch, eine Veränderungseinschätzung geben zu müssen, zu Antworten im Sinne „einer Tendenz zur sozialen Erwünschtheit ... oder eines eigenen Wunsches, welcher der Realität aber nicht entspricht“ (Langenmayr und Kosfelder 1998, S. 101), greifen. Darüber hinaus besteht bei jeder retrospektiven Einschätzung ja auch das Risiko, der Neigung zu erliegen, „den Erfolg der Therapie günstig einzuschätzen, damit ... (das) ... Selbstbild nicht leidet“ (Langenmayr und Kosfelder 1998, S. 104) oder es besteht die Gefahr, Übertragungswirkungen im Sinne „einer mangelnden Ablösung“ (Langenmayr und Kosfelder, 1998, S. 101) zu erliegen.

d) „Schließlich ist das Umfeld der Studie nicht durch eine artifizielle Versuchsanordnung charakterisiert, sondern es handelt sich um eine repräsentative Stichprobe aller ambulanten Behandlungen, die bundesweit in einem bestimmten Zeitraum durchgeführt wurden“ (Heinzel u.a., 1996, S.2). Langenmayr und Kosfelder sehen in der Frage der repräsentativen Stichprobengewinnung die Grundlage dafür, dass die aus den Daten gezogenen Schlussfolgerungen verallgemeinerungsfähig sind und sprechen in diesem Zusammenhang von der „externen Validität“ (1998, S. 104) der Ergebnisse. Was unsere Studie anbelangt, so können wir nicht garantieren, so aber doch mit einigem Recht vermuten, dass die untersuchten Stichproben repräsentativ sind. Bei unserer Untersuchungsgruppe handelt es sich nicht um eine artifizielle Patientengruppe zu Untersuchungszwecken, auch nicht um ‚gesunde Kranke‘, wie sie z.B. in universitären Beratungsstellen für studienmüde Studenten vorzufinden sind, sondern es handelt sich um tatsächliche über die Krankenkassen finanzierte Therapien, die bundesweit im Zeitraum 1992 – 1998 von individualpsychologischen Psychotherapeuten durchgeführt wurden. Die Zufälligkeit der Stichprobengewinnung erfuhr allerdings eine Einschränkung, nämlich an der Bereitschaft der Therapeuten, ihre Patienten überhaupt zur Mitwirkung an der Studie gewinnen zu wollen. Unsere Stichprobe bezieht sich also nur auf Patienten von Therapeuten, die zur Mitwirkung bereit waren. Patienten, die bei Therapeuten in Behandlung waren, die an der Studie nicht mitwirken wollten, hatten keinen Zugang zur Studie. Es wäre zu diskutieren, ob die Variable „mitwirkungsbereite Therapeuten“ bzw. „nicht mitwirkungsbereite Therapeuten“ die Repräsentativität der Stichprobe beeinflusst hat und diese Variable darüber hinaus auch bereits eine den Therapieerfolg beeinflussende Wirkgröße darstellt. Hierzu können wir im Rahmen unserer Studie keine Aussagen machen.

e) Zum Problem der Schaffung einer quasi-experimentellen Kontrollgruppe sagen Heinzel u.a.: „Es scheint vor allem davon abzuhängen, ob man für eine bestimmte Zahl von behandelten Patienten eine passende Kontrollgruppe bilden kann“ (Heinzel u.a., 1996, S. 7). Langenmayr und Kosfelder (1998) sehen in der Frage, ob sich Untersuchungs- und Kontrollgruppe „schon zu Beginn der Behandlung unterscheiden“, einen „Störfaktor der externen Validität“ (1998, S. 104).

Bei unserer Kontrollgruppe, die rein zufällig zustande gekommen ist, wussten wir zunächst nicht, ob es sich um eine „passende“ Kontrollgruppe handelte, die in ihrer Zusammensetzung hinsichtlich bestimmter Merkmale der Untersuchungsgruppe vergleichbar war. Dies konnten wir erst im Nachhinein mit Hilfe statistischer Prüfverfahren erkunden (vgl. unten Punkt 10.3.2.1). Diese Prüfverfahren bestätigten jedoch, dass sich Untersuchungsgruppe und Kontrollgruppe hinsichtlich der geprüften Merkmale weitgehend ähnlich waren, so dass wir unsere Kontrollgruppe als „passende“ Kontrollgruppe ansehen können. Langenmayr und Kosfelder machen jedoch einschränkend auf das Problem aufmerksam, dass die Prüfung der Strukturgleichheit von Untersuchungs- und Kontrollgruppe immer nur mit Vorbehalt behauptet werden kann, weil die „relativ große Zahl der zu berücksichtigenden Parallelisierungsmerkmale (Alter, Geschlecht, Störungsdauer und -schwere, etc.)“ (1998, S. 105) jeder Strukturgleichheitsprüfung natürliche Grenzen setzen.

Zum Problemkreis ‚Kontrollgruppe‘ schreibt Grawe, dass die Daten für die „Kontrollgruppe ... über den gleichen Zeitraum wie bei den psychotherapeutisch behandelten Patienten erhoben werden“ (Grawe, 1992, S. 134) müssen, um Vergleichbarkeit zu erzielen. Dieses hohe Idealmaß konnten wir allerdings nicht erfüllen,

denn kein ‚real kranker Patient‘ ist aus Forschungsgründen bereit, u.U. mehrere Jahre lang auf eine Behandlung zu warten, nur um sich aus wissenschaftlichen Forschungszwecken einer Kontrollgruppe zur Verfügung zu stellen. Umgekehrt wollten wir aber auch nicht auf die Kontrollgruppe verzichten. Wir haben uns deshalb zu einem Kompromiss, - angelehnt an den real möglichen und machbaren Bedingungen -, durchgerungen. Wohl wissend, dass viele Kollegen Wartelisten führen, konnten wir keine gravierenden ethischen Einwände erkennen, die uns hätten abhalten können, diese so wartenden Patienten zu fragen, ob sie bereit wären, an unserer Untersuchung als Probanden mitzuwirken. Also wohl gemerkt: Wir haben unsere Kollegen nicht gebeten, Patienten absichtlich ‚warten zu lassen, um Probanden für unsere Kontrollgruppe sein zu können‘, sondern wir haben unsere Kollegen gebeten, die Patienten, die sie infolge ihrer ganz natürlichen Praxisüberlastung sonst abgewiesen hätten oder ggf. auf einer Warteliste führten, nach ihrer eventuell vorhandenen Mitwirkungsbereitschaft im Rahmen unserer Untersuchung zu befragen.

Wir gehen auch davon aus, dass kein Patient länger als nötig, also sozusagen ‚künstlich‘ auf der Warteliste ‚geblieben ist‘, nur um das „N“ in unserer ‚Kontrollgruppe‘ zu erhöhen. Als Beleg hierfür mag gelten, dass die Kontrollgruppe bereits zum Messzeitpunkt t2 (also einem halben Jahr nach der Erstbefragung) und nach einem weiteren halben Jahr, also zum Messzeitpunkt t3, noch mehr zusammengeschmolzen war. Dass der antwortende Patient zwischenzeitlich nicht mehr auf der Warteliste stand, sondern die Therapie bereits mit einer unterschiedlich großen Anzahl von bereits stattgefundenen Sitzungen begonnen hatte, erfuhren wir in allen Fällen immer erst durch den Rücklauf unserer verschickten Testunterlagen. Die Begrenzung der Kontrollgruppe auf einen Zeitraum von einem Jahr begründen wir wie folgt: Wenn es sich zeigt, dass die Untersuchungsgruppe nach einem Jahr Psychotherapie signifikant höhere Zustandsverbesserungen als die Kontrollgruppe aufzuweisen hat, dann erscheint uns dies ein hinreichendes Ergebnis zu sein, um den prinzipiellen Nachweis der Wirksamkeit von Psychotherapie erbracht zu haben. Ob eine ‚Testverlängerung‘ auf mehrere Jahre dieses Ergebnis dann noch um einige Signifikanzpunkte verbessert hätte, soll uns nicht interessieren.

f) Heinzel u.a. verweisen auf das dem empiristischen Denken entstammende Problem, „dass die Qualität der Einschätzungen seitens der Therapeuten, Forscher und Patienten von deren Motivation abhängt“ (Heinzel u.a., 1996, S. 12). Was unsere Untersuchung anbelangt, so kann dieser Einwand, - zumindest was die Therapeuten und Forscher anbelangt -, entkräftet werden, da Therapeuten und Forscher ja gar nicht erst befragt wurden, so dass deren Motivation zumindest in direkter Weise keinen Einfluss auf das Untersuchungsergebnis genommen haben kann. Bleibt zu diskutieren, ob die Motivation der Therapeuten auf indirektem Weg Einfluss auf das Ergebnis genommen hat, z.B. dadurch, dass sie aus dem Wissen um die ihren Patienten geltende Forschungsaufmerksamkeit heraus höher therapiemotiviert waren als in anderen Fällen. Dann wären die festgestellten Therapieeffekte Ausdruck dieser besonderen forschungsinduzierten Motivation (vgl. Hawthorne-Effekt, Langenmayr und Kosfelder, 1998, S. 105). Wenn man unterstellt, dass die Patienten ebenfalls von dieser forschungsinduzierten Motivation beflügelt waren, könnte man deren Veränderungen als weitestgehend forschungsinduziert ansehen. Diese Sichtweise mag zunächst erschrecken, bei näherem Hinsehen jedoch eher erfreuen. Denn was wäre an diesem Ergebnis so verwerflich (?), - sagt es doch nur, dass

motivierte Therapeuten und motivierte Patienten zu besseren Therapieergebnissen als schlecht motivierte Therapeuten und/oder schlecht motivierte Patienten gelangen, was ja nicht verwunderlich ist. Es wäre also mit solch einem Ergebnis belegt, dass Psychotherapie um so erfolgreicher sein kann, je höher die Motivation der Beteiligten ist.

Unabhängig hiervon kann der Einwand, dass die ggf. nachgewiesene positive Wirkung nur auf das Wissen um die Therapie zurückzuführen ist, mit Hilfe der Kontrollgruppe überprüft werden, denn auch die Patienten der Kontrollgruppe wissen ja, dass sie an einer wissenschaftlichen Untersuchung teilnehmen. Wenn dem aber so ist, dann sind Untersuchungs- und Kontrollgruppe in dieser Variablen ‚Wissen um die Mitwirkung an einem Forschungsprojekt‘ gleich, der in diesem Falle vorliegende ‚Validitätsfehler unserer Messung‘ wäre ein beide Gruppen betreffender systematischer Fehler, so dass eventuelle Unterschiede zwischen beiden Gruppen nicht mehr auf eine in diesem Punkt vermutete unterschiedliche Motivationsquelle zurück zu führen sind, sondern dann eben doch nur das Ergebnis der unterschiedlichen Therapievariablen sein kann.

Langenmayr und Kosfelder machen auf die Möglichkeit umgekehrt wirkender Effekte aufmerksam, dass nämlich z.B. „umfangreiche Testbatterien die Motivation zum sorgfältigen Ausfüllen von Fragebögen stark senken und somit zu einer Verzerrung der Ergebnisse führen“ (1998, S. 104) können.

Wenn mit einer Motivationssenkung zu rechnen ist, bedeutet dies aber doch lediglich, dass sich die oben diskutierte forschungsinduzierte Motivation nach einiger Zeit verschleißt und die Patienten eher unüberlegter, damit aber eher wahrheitsgerechter antworten, was doch eher für die Validität unserer Ergebnisse und nicht dagegen spricht.

Es bleibt eine weitere zugespitzte Frage, nämlich ob die Patienten in ihrem Antwortverhalten hinsichtlich ihres jeweilig aktuellen Befindens von ihrer augenblicklichen Tagesmotivation beeinflusst wurden. Natürlich muss man dies bejahen, doch kann man davon ausgehen, dass die Menge der individuell zu beantwortenden Items in GBB, MMPI(K) und VEV eine etwa inhaltlich ausgerichtete Anfangsmotivation im Einzelfall sehr schnell zunichte gemacht hat, weshalb man doch von einem überwiegend eher befindlichkeitsorientierten Antwortverhalten ausgehen kann. Im übrigen kann davon ausgegangen werden, dass sich etwaige unterschiedliche Individualmotivationen über das doch relative hohe „N“ der Patienten ausgeglichen/nivelliert haben müssten.

Bleibt die Frage, ob die Motivation der Forscher Einfluss auf das Untersuchungsergebnis genommen hat. Wir können dies mit dem Hinweis verneinen, dass die Forscher alles getan haben, - ja sogar den Forschungsansatz der Kritiker gewählt haben -, um alle Möglichkeiten zu eröffnen, einem gewünschten Ergebnis Falsifikationsmöglichkeiten bereit zu stellen.

g) Heinzel u.a. verweisen des weiteren auf den möglichen methodischen Einwand, „dass die Anwesenheit des ‚Evaluators‘ den Patienten beeinflusst“ (Heinzel u.a., 1996, S.12). In unserem Falle waren die Evaluatoren bei der Beantwortung der Items nicht nur nicht anwesend, sondern die Evaluatoren und die Patienten kannten sich nicht, kamen nie zusammen und hatten lediglich einen einseitig verlaufenden Schriftverkehr derart miteinander, dass die Befindlichkeitsfragebögen von den Forschern an die Therapeuten verschickt, von diesen zur häuslichen Bearbeitung an die Patienten weitergeleitet und von diesen dann wiederum ano-

nym an die Forschungsgruppe zurückgesandt wurden. Die Patienten füllten die Fragebögen also ohne Anwesenheit eines u.U. interessengebundenen Forscher-, Evaluator- oder Therapeuteneinflusses aus.

h) Heinzl u.a. verweisen auf das für die empirische Wissenschaftsauffassung ungelöste methodische Problem, „dass die Art des Messinstruments, Fragebogens usw., einen Einfluss auf die Einschätzungen hat“ (Heinzl u.a., 1996, S.12). Dieses Argument greift jedoch nur bei der Bewertung einer singulären Einzelmessung, nicht jedoch wenn es darum geht, einen Messunterschied zwischen verschiedenen mit dem gleichen Messinstrument durchgeführten Einzelmessungen zu ermitteln. In diesem Fall schlägt sich nämlich der messinstrumentbedingte Einfluss in allen Wiederholungsmessungen nieder, - wodurch dieser „Messfehler“ zu einem „systematischen Messfehler“ avanciert. Dies aber heißt, dass für unsere Untersuchung, die ja hauptsächlich mit Wiederholungsmessungen arbeitet, davon auszugehen ist, dass in jede Einzelmessung zwar systematische Messfehler eingegangen sind, dass aber eben gerade durch das ‚Systematische‘ des Messfehlers der Unterschied/Abstand zwischen den verglichenen Einzelmessungen nicht tangiert wird.

i) Heinzls u.a. Warnung vor dem Einwand, „...dass die gemessenen Therapieeffekte unterschiedlich ausfallen, je nachdem, aus welcher Quelle sie stammen“ (Heinzl u.a., 1996, S. 12), ist uns inhaltlich unverständlich. Was meinen Heinzl u.a. mit „Quelle“? Insofern sie damit Patient/Therapeut/Forscher verstehen, so haben wir hierauf oben bereits geantwortet. Insofern er darunter die wissenschaftsorientierte und/oder interessenbezogene Grundsatzposition des Forschers meint, so könnte damit der mögliche empiristische Vorwurf gegen unserer Untersuchung gemeint sein, dass die beteiligten Forscher, - selbst aus dem analytischen Lager kommend -, unreflektiert und unkontrolliert das Forschungsinteresse in die Untersuchung mit einbrachten, die Effektivität von Langzeitanalysen nachweisen zu wollen, weswegen es dann natürlich kein Wunder sei, dass sie das Forschungssetting so beeinflussten, dass die derart interessenbezogene Forschung auch die gewünschten Resultate liefere. Dem können wir, - wie oben bereits ausgeführt -, entgegenhalten, dass die durchführenden Forscher es sich in diesem Falle besonders schwer gemacht haben, indem sie nämlich die wissenschaftstheoretischen und forschungslogischen Kriterien ihrer „Kritiker“ nicht nur halbherzig berücksichtigt, sondern sich die Einhaltung dieser Kriterien geradezu auf die Fahnen geschrieben haben und damit hinsichtlich eines zugegebenermaßen gewünschten Resultats allen „Falsifikationsmöglichkeiten“ bis hin zur Auswahl der Messverfahren zugearbeitet haben.

j) Heinzls u.a. Warnung vor dem Einwand einer „zu kleinen Stichprobe.... (Nach Bergin und Garfield 1994 ist die durchschnittliche Stichprobengröße bei aller ihnen verfügbaren Untersuchungen zwölf!)“ (Heinzl u.a., 1996, S. 13) können wir entgegen halten, dass wir in unserer Untersuchung ein sehr viel größeres „N“ erzielten: Zum Zeitpunkt t1 bestand die Untersuchungsgruppe aus  $n = 144$  und die Kontrollgruppe aus  $n = 50$ . Zum Zeitpunkt t3 hatte die UG noch 110, die KG noch 39 Mitglieder. Die KG wurde ab diesem Zeitpunkt nicht weiter fortgeführt, die UG fiel von  $n = 85$  zum Messzeitpunkt t4 auf  $n = 73$  zum Messzeitpunkt t5 (drei Jahre nach Therapiebeginn) ab. *Langenmayr* und *Kosfelder* (1998) machen auf die Gefahren sowohl einer zu kleinen wie auch einer zu großen Stichprobe aufmerksam. Eine zu große Stichprobe produziere u.U. statistische Signifi-

kanzen, die in der Praxis irrelevant seien, während eine zu kleine Stichprobe das Risiko der unbemerkten Unkenntlichkeit vorhandener Signifikanzen in sich berge. Sie geben eine Stichprobengröße von „zwischen 30 und 50 KlientInnen“ als Richtgröße an, die „den Erfordernissen am ehesten gerecht“ (1998, S. 107) werde. Diese Richtgröße erschien uns unter dem Aspekt, dass wir planten, für unsere Langzeitstudie über mehr als drei Jahre Daten zu erheben, als zu gering angesetzt, weil wir einen normalen Patientenschwund mit einkalkulieren mussten, weshalb wir sozusagen aus einer Vorsorgehaltung heraus mit der o.g. wesentlich größeren Anfangsstichprobe in unsere Studie starteten.

k) Heinzl u.a. warnen weiterhin vor dem möglichen empiristischen Einwand, die „Patientenauswahl entspricht nicht den praktischen Gegebenheiten (freiwillige Versuchspersonen, ‚Patienten‘ mit überwiegend leichten Störungen)“ (Heinzl u.a., 1996, S.13). Auch dieses Bedenken können wir für unsere Untersuchung zurückweisen. Die Patienten entstammen den tatsächlichen Gegebenheiten real existierender Praxen (naturalistische Untersuchung). Man kann uns natürlich vorhalten, dass wir nicht kontrolliert haben, ob die mitwirkenden Therapeuten u.U. nur solche Patienten auf ihre eventuelle Mitwirkungsbereitschaft angesprochen haben, die ihnen als „leichte Patienten“ erschienen haben mögen. Dieser Einwand ist für unsere Studie zumindest aus der Anfangssicht noch nicht vorliegender Prüfmöglichkeiten zur Schwere der Erkrankungen theoretisch richtig.

Zu Beginn wussten wir lediglich, dass für alle Patienten gemäß den Grundsätzen der „Richtlinienpsychotherapie“ der Krankenkassen ein Gutachterverfahren durchlaufen worden war. Ob es sich hierbei um ‚gesunde Kranke‘ oder um ‚kranke Kranke‘ handelte, konnten wir zu diesem Zeitpunkt nicht wissen. Erst im Nachhinein, also nach Vorliegen der Messergebnisse zum Messzeitpunkt t1, konnten wir die in den klinischen Skalen gemessenen Werte unserer Patienten mit den Werten der in der testpsychologischen Literatur veröffentlichten Eichstichproben vergleichen (hierzu vgl. unten Punkt 5.4). Durch diese Prüfungen ergab sich dann allerdings, dass es sich bei unseren Patienten tatsächlich um ‚kranke Kranke‘ und nicht um ‚gesunde Kranke‘ handelte (vgl. Kapitel 10.3.4).

l) Ein weiterer Einwand gegen die bisherigen psychoanalytischen Wirksamkeitsstudien lautet: „Die Therapeuten sind nicht erfahren genug. (Nach Parloff (1982) waren die meisten Therapeuten Doktoranden mit einer mittleren therapeutischen Erfahrung von 3 Jahren!).“ (Heinzl u.a., 1996, S.13) Auch diesen Einwand können wir für unsere Untersuchung entkräften. Bei unseren Therapeuten handelte es sich nicht um wissenschaftlich interessierte ‚Hochschultherapeuten mit Assistentenstatus‘, sondern es handelte sich ausnahmslos entweder um niedergelassene oder in Kliniken angestellte Psychotherapeuten mit einer durchschnittlichen Berufspraxis von im Schnitt 6,2 Jahren (zum Ausbildungsstand der mitwirkenden Therapeuten vgl. unten, Punkt 5.3).

m) Ein weiterer von Heinzl u.a. referierter möglicher empiristischer Einwand: „Die Erfolgskriterien sind nicht adäquat bzw. uneinheitlich.“ (Heinzl u.a., 1996, S.13) Langenmayr und Kosfelder (1998) unterscheiden für die Evaluation grundsätzlich zwei mögliche Erfolgskriterien, nämlich „einerseits die Veränderung gegenüber dem Therapiebeginn und/oder andererseits die Annäherung an vorab definierte Ziele“, z.B. das „Absinken ...



unter eine klinisch relevante Schwelle“ (1998, S.102). Wir haben in unserer Studie beide Kriterien zu berücksichtigen gesucht, nämlich zum einen der Vergleich von Behandlungsanfang und Behandlungsende, zum anderen aber auch den Vergleich der gemessenen Kennwerte mit externen Kennwerten z.B. der gesunden und kranken Eichpopulationen der angewendeten Messverfahren.

Als Erfolgskriterium der Prä-Post-Messungen weisen wir das Erreichen statistischer Signifikanzkriterien aus. Als Erfolgskriterien im externen Vergleich beziehen wir uns auf das Ausmaß der Annäherung an die externen Norm- und Eichwerte, - hauptsächlich des GBB und des MMPI(K). Beide Verfahren stellen in Hinblick auf unsere Studie unabhängige Messinstrumente mit ausgewiesenen eigenen Eichwerten sowohl für die ‚gesunde Normalbevölkerung‘ wie auch für ‚kranke Menschen‘ dar. Stellt sich die Frage, ob es sich hierbei um adäquate Messinstrumente hinsichtlich des Messgegenstandes „Therapieerfolg“ handelt. Diese Frage berührt die Problemkreise der Validität und Reliabilität unserer Messinstrumente, die wir weiter unten diskutieren (vgl. Kapitel 5.12.1)

n) Heinzel u.a. verweisen auf das Problem der „Uneinheitliche(n) Länge der Therapien. (Nach Henry et al. 1994 ist die durchschnittliche Gesamtzahl der Sitzungen der untersuchten Therapien zwischen 20 und 150, die Mehrheit unter 50)“ (Heinzel u.a., 1996, S.13). Vermutlich ist der empiristische Einwand gemeint, dass die untersuchten Therapien innerhalb einer Untersuchung nie gleich lang sind.

Die von uns untersuchten Therapien sind natürlich auch nicht alle gleich lang, - das liegt in der Natur von Therapien und lässt sich nun einmal nicht ändern, solange Therapien nicht ‚am Patienten vollstreckt‘ werden. Bei den von uns untersuchten Therapien handelt es sich um Langzeittherapien. Zum Messzeitpunkt t3 (ein Jahr nach Therapiebeginn) waren noch 106 Therapien mit einem Durchschnittswert von 70 Sitzungen pro Therapiefall an der Untersuchung beteiligt. Zum Messzeitpunkt t5 (drei Jahre nach Therapiebeginn) waren immerhin noch 72 bis dahin nicht abgeschlossene Therapien mit einem Durchschnittswert von bis dahin erreichten 198 Sitzungen pro Therapie in der Untersuchung. Wir leiten aus diesen Werten die Berechtigung ab, unsere Ergebnisse als Aussagen über wirkliche Langzeitpsychotherapien deklarieren zu können.

o) Heinzel u.a. referieren die empiristische Kritik an den bislang bekannten Wirkungsforschungen, dass nämlich „Nachfolgeuntersuchungen fehlen“ (Heinzel u.a., 1996, S.13). Wir vermuten, dass Heinzel mit „Nachfolgeuntersuchungen“ katamnestische Untersuchungen meint und merken an, dass eine katamnestische Nachbefragung ganz normaler Bestandteil unserer Untersuchung ist (vgl. unten Kap. 11).

p) Heinzel u.a. verweisen auf eine erkenntnistheoretische Grundbedingung: „Schon seit Heisenberg wissen wir, dass die Beobachtung eines Vorgangs diesen grundsätzlich beeinflusst.“ (Heinzel u.a., 1996, S. 14) Wir haben zwar keine ‚teilnehmende Beobachtung‘ durchgeführt, gleichwohl trifft dieser Einwand natürlich auch für unsere Untersuchung zu und wir müssen eingestehen, dass wir diese Beeinflussung nicht zu kontrollieren oder in ihrer Größenordnung zu erfassen suchten. Wir konnten nämlich nicht ausschließen, - dies haben wir bereits oben diskutiert - , dass das Wissen der Patienten, als Teilnehmer an einer Untersuchungs- oder an einer Kontrollgruppe über das Medium Fragebögen ‚beobachtet‘ zu werden, das Therapiegeschehen, und

damit natürlich auch das Antwortverhalten beim Ausfüllen unserer Messfragebögen mit beeinflusst hat, - und wir können auch keine Aussage über Art und Umfang dieser Beeinflussung treffen. Wir können nur, - wie bereits erwähnt -, mit einigem Recht vermuten, dass diese Beeinflussung des „Messgegenstandes“ eine systematische gewesen sein muss, denn sie hat ja Untersuchungs- und Kontrollgruppe im gleichen Maß betroffen, - woraus aber wieder zu folgern ist, dass eine unterschiedliche Entwicklung von Untersuchungs- und Kontrollgruppe nicht Ausdruck dieses „systematischen Fehlers“, sondern nur Ausdruck der unabhängigen Variablen „Therapieeinfluss“ sein kann.

q) Heinzel u.a. weisen darauf hin, „dass (u.a. bei Persönlichkeitsstörungen) das Gegenteil geschieht, nämlich die Verleugnung einer Besserung“ (Heinzel u.a., 1996, S. 12), was den Therapieerfolg schwerer messbar macht. Dies mag stimmen oder nicht, - jedenfalls wäre dies ja nur von Bedeutung für Untersuchungen, die keinen oder einen nur marginal messbaren Therapieerfolg nachweisen können. Denn nur in diesem Fall hat es ja Sinn, die Überlegung anzustellen, ob es u.U. einen Therapieerfolg gegeben hat, der jedoch auf Grund der „Verleugnung der Besserung“ nicht messbar ist. Für Untersuchungen, die einen Therapieerfolg signifikant nachweisen können, ist es letztlich irrelevant, ob der Wirksamkeitsnachweis nun noch etwas deutlicher hätte ausfallen können oder nicht. Glücklicherweise gehört unsere Untersuchung vom Ergebnis her zu der zweiten Art.

r) Heinzels u.a. vorgenannte Überlegung, dass Therapieerfolg u.U. nicht messbar ist, weil die Besserung verleugnet wurde, macht auf ein weiteres Grundproblem aufmerksam, dass es nämlich unterschiedliche Sorten von Erfolg gibt, „Erfolg“ also nicht eindeutig definierbar ist. Langenmayr und Kosfelder verweisen in diesem Zusammenhang auf die „paradoxe Situation...“, dass ein negatives Urteil zu einem bestimmten Zeitpunkt vor Abschluss der Therapie aus therapeutischer Perspektive günstig zu bewerten ist“ (1998, S. 101). So kann ein Erfolg ja u.U. auch darin bestehen, dass es dem Patienten gelingt, eine aggressive Trennung vom Therapeuten durchzuführen. Formal liefe dieser Patient unter ‚Abbrecher‘.

Tatsächlich aber kann es ja durchaus sein, dass die Therapie bei einem solchen Patienten über das Zustandbringen einer erstmalig gelungenen aggressiven Trennung langfristig große Veränderungen freisetzt. Man kann diesem Argument natürlich aus einer theoretisch besetzten Grundhaltung heraus entgegenhalten, dass eine ‚aggressive Trennung‘ niemals ein Therapieerfolg sein könne, weil die negative Übertragung nicht aufgelöst sei. Dagegen steht, dass Seelisches immer Zeit braucht und sich nur in übergreifenden Wirkungseinheiten entwickelt, die allerdings manchmal an der Theorie-Realität vorgegebener Zeiteinheiten vorbei gehen. In diesem Zusammenhang verweisen wir auf eine vom Grundansatz her ähnliche ‚umwertende‘ Überlegung bei Hampe: „Voraussetzung einer positiven Bewertung des psychoanalytischen Prozesses scheint die Tatsache von Leidensminderung zu sein. Eine ethische Rechtfertigung, dass Leidensminderung immer gut ist, findet nicht statt.“ (Hampe, 2001, S.333). Wir möchten diesen Gedanken hier auch einmal auf den Begriff der ‚gelösten bzw. ungelösten Übertragung‘ angewendet wissen. Unter Hinweis auf andere Autoren referieren auch Heuft u.a. von Fällen, in denen „die Übertragung mit der Behandlungszeit zunahm und ‚hoch‘ blieb,

unabhängig vom Outcome. Dem entsprechen Befunde ..., die durch ihre ‚Katamneseanalysen‘ auch bei erfolgreich analysierten Personen eine Wiederbelebung der Übertragungsneurose registrieren konnten“ (Heuft u.a., 1996, S. 354). Für unsere Studie sind solche Überlegungen jedoch irrelevant, weil sie bereits einem „nicht mehr positivistischen“, sondern eher „analytischen Denken“ entspringen. Da wir uns aber entschlossen haben, in dieser Studie den empiristischen Ansatz möglichst streng durchzuhalten, sind für uns alle „Antwortkreuzchen“ der Patienten unhinterfragbare (positive) Tatsachenentscheidungen, in denen subjektiv empfundenen Zustände zu objektiven Daten umgewandelt worden sind.

Die bereits mehrfach erwähnten Autoren Langenmayr und Kosfelder (1998) haben im Nachgang zu Grawe die Standards der zeitgemäßen Psychotherapieforschung über Heinzel u.a. hinausgehend systematisch zusammengestellt und damit eine allgemeine Checkliste der Evaluation geliefert, an der auch unserer Studie gemessen werden kann. Langenmayr u.a. nennen acht Leitfragen/Kriterien/Standards, die explizit zu berücksichtigen und einzuhalten letztlich keine Psychotherapiestudie umhin kann, ohne angesichts der zeitgenössischen Wissenschaftslehre ihre wissenschaftliche Relevanz zu verlieren. Diese Leitfragen betreffen a) die explizite Nennung der Zielsetzung einer Studie, b) eine bewusste Entscheidung über die Ebene der zu messenden Therapieeffekte, c) die Festlegung der Operationalisierung der vorgesehenen Messungen, d) die vorherige Nennung der Kriterien für den „Therapieerfolg“, e) die Planung methodologischer Vorkehrungen zur Kontrolle und Isolierung aller beteiligten Variablen, f) die Nennung von Maßnahmen zur Absicherung der Aussagegültigkeit und der Generalisierbarkeit der Forschungsergebnisse, g) die Einhaltung eines Studiendesigns mit Prä-Post-Messungen und katamnestischer Kontrolle sowie h) die Berücksichtigung statistischer Kennwerte bei der Auswertung der Daten.

Unsere Untersuchung sucht alle diese Kriterien von Langenmayr und Kosfelder wie folgt zu erfüllen: a) Die Zielsetzung unserer Studie besteht darin, die Wirksamkeit der Individualpsychologisch-Psychoanalytischen Psychotherapie zu belegen (Outcome-Forschung) und einige ausgewählte Wirkfaktoren zu überprüfen, b) die Erfolgsmessung soll mit störungsspezifischen Maßen auf der Ebene der Symptome und Beschwerden erfolgen, c) die Erfassung (Operationalisierung) der zu messenden symptomatischen Krankheits-Messwerte erfolgt im Kernbereich der Studie mittels der symptombezogenen Messverfahren des MMPI(K), des GBB und des VEV, - im katamnestischen Teil mittels des HAQ und der Intrex-Kurzform (vgl. Kap. 11.2.3.1-2), d) als Erfolgskriterien werden Veränderungen zwischen Vorher- und Nachher-Messungen sowie unter besonderer Berücksichtigung des Konsens hinsichtlich des VEV auch rückwirkende Einschätzungen der Klienten herangezogen, e) um die gemessenen Veränderungen auf die Therapievariable zurückführen zu können, wird ein Kontrollgruppendesign gebildet und die Strukturgleichheit von Untersuchungs- und Kontrollgruppe geprüft, f) zur Absicherung der Aussagegültigkeit und der Generalisierbarkeit der Ergebnisse wird ein naturalistisches Studiendesign mit repräsentativer Zufallsauswahl der beteiligten Patienten und anonymisierter Befragungstechnik der Patienten unter Ausschaltung der Therapeuten favorisiert, g) um den Verlauf der Veränderungen zu verfolgen, werden fünf Messzeitpunkte vor und während der Behandlung sowie eine katamnestische Nachbefragung ein Jahr nach Therapie-Ende geplant und durchgeführt, h) die Datensammlung wird hinsichtlich der Erreichung statistisch ausreichend großer Stichproben zusammengetragen und mittels anerkannter

statistischer Prüfverfahren auf das Erreichen aussagekräftiger Kennwerte (Signifikanzniveaus) und Effektstärken hin untersucht.

Wir denken, damit allen Kriterien der empirischen Sozialforschung ausreichend zu genügen und für die Gültigkeit, Verallgemeinerbarkeit und Übertragbarkeit unserer Ergebnisse auf das Therapieverfahren der „Individualpsychologisch-Psychoanalytischen Psychotherapie“ im Ganzen hinreichend Sorge getragen zu haben.

## 5. Vorentscheidungen

### 5.1 Begrenzung auf „Individualpsychologisch-Psychoanalytische Psychotherapie“

Wir standen vor der Entscheidung, ob sich unsere Untersuchung auf alle psychoanalytischen Schulrichtungen erstrecken oder auf den Bereich der „Individualpsychologisch-Psychoanalytische Psychotherapie“ begrenzen sollte. Da Grawe im „Forschungsgutachten“ den psychoanalytischen Therapien der Freudianischen Schulen jedoch bereits eine Wirksamkeit zugesprochen, gleichwohl der sogenannten „Individualtherapie“ aber vorenthalten hatte, entschlossen wir uns, unsere Untersuchung auf diese Schulrichtung einzugrenzen. Dabei sind wir der „adherence“-Problematik, also der Frage, ob die mitwirkenden Therapeuten die von ihnen ausbildungsmäßig angebotene und von uns vermeintlich untersuchte Therapieform, hier „Individualpsychologisch-Psychoanalytische Psychotherapie“, auch ‚fachgerecht‘ angewendet haben, so dass wir auch wirklich mit Recht sagen können, „Individualpsychologisch-Psychoanalytische Psychotherapie“ untersucht zu haben, nicht weiter nachgegangen. Dies deshalb, weil uns diese Frage Rahmen eines empirisch-positivistischen Grundverständnisses ja gar nicht interessieren muss. Denn was ist „Individualpsychologisch-Psychoanalytische Psychotherapie“? Welche Grundsätze unterscheidet die „Individualpsychologisch-Psychoanalytische Psychotherapie“ von anderen Therapieformen?

Lassen sich diese, - wenn es sie denn gibt -, besonderen Kennzeichen einer „Individualpsychologisch-Psychoanalytischen Psychotherapie“ überhaupt intersubjektiv objektivieren, operationalisieren und manualisieren? Könnten wir diese Variable konstant halten? Zweifel über Zweifel. Stattdessen gehen wir in Anlehnung an die bekannte empirisch-positivistische Definition „Intelligenz ist das, was der Intelligenztest misst“ hier sehr empirisch-positivistisch vor: „Individualpsychologisch-Psychoanalytische Psychotherapie“ ist das, was an anerkannten individualpsychologischen Instituten ausgebildete individualpsychologische Psychotherapeuten mit ihren Patienten tun.“

### 5.2 Trennung von Therapeut und Untersucher

Wir, die Durchführenden der Untersuchung, hatten zunächst vor, unsere eigenen Patienten in die Untersuchung mit einzubringen, - schon alleine aus dem Grunde, um ein für die statistischen Prüfverfahren hinreichend großes „N“ zu erzielen. Auch hätte uns dies ein Teil der Mühe erspart, unsere Kollegen zur Mitwirkung an der Untersuchung zu motivieren. Aus methodologischen Gründen entschlossen wir uns jedoch dazu, hierauf zu verzichten, auch wenn wir uns aus den eigenen Reihen der Kritik ausgesetzt sehen, dem lägen auf Forscherseite die gleichen Ängste und Widerstände zugrunde wie sie auf Seiten der sich verweigernden Therapeuten zu finden seien. Wir waren zwar zunächst geneigt, eigene Patienten mit in die Untersuchung

hinein zu nehmen, davon abgehalten hat uns dann jedoch die Überlegung, dass wir durch die Vermischung der Rollen von „Forscher und Therapeut“ einem Grundaxiom der „Methodik der empirischen Sozialforschung“ zuwider gehandelt und unsere Ergebnisse damit sofort der methodologischen Kritik des Empirismus ausgesetzt hätten, das Ergebnis nämlich unbewusst im Sinne der gewünschten Ergebnisse verfälscht zu haben, z.B. allein schon durch die Manipulationsmöglichkeit in der Phase der Patientenauswahl, möglichst viele ‚leichte‘ und somit erfolgversprechende Fälle in die Untersuchung einzuschleusen.

### 5.3 Ausbildungsstand der Therapeuten

Um sicher zu sein, auch tatsächlich nur fachlich fundiert durchgeführte Therapien zu untersuchen, suchten wir nach einem Standard. Diesen Standard (Operationalisierung) fanden wir darin, dass wir nur Psychoanalytiker mit einer abgeschlossenen Ausbildung an einem anerkannten Ausbildungsinstitut der DGIP mit ihren Patienten in die Untersuchung hinein nahmen. Wir ließen uns also von jedem Therapeuten mitteilen, in welchem Jahr und an welchem Ausbildungsinstitut er die Ausbildung abgeschlossen hat. Insgesamt stellten wir bei Beginn der Untersuchung, also zum Messzeitpunkt t1, bezogen auf 54 zum damaligen Zeitpunkt mitwirkungsbereite Therapeuten ein Mittel der beruflichen Tätigkeit von 6,2 Berufsjahren fest. 19 Therapeuten wiesen weniger als 4 Berufsjahre auf, 25 Therapeuten lagen in der Gruppe von vier bis neun Berufsjahren, neun Therapeuten lagen in der Gruppe zehn bis neunzehn Berufsjahren, eine Therapeutin erreichte 26 Berufsjahre. Wir sehen darin den Nachweis, dass die Untersuchung sich im wesentlichen auf berufserfahrene Therapeuten stützt.

### 5.4 Krankheitsnachweis

Es gibt im Bereich der Psychotherapie ‚gesunde Kranke‘ und ‚kranke Kranke‘. Für unsere Untersuchung wollten wir sicher sein, dass wir nicht ‚gesunde Kranke‘ untersuchen. Dies schien uns dadurch hinreichend gesichert, dass wir sicherstellten, dass es sich bei allen Patienten nur um solche handelte, deren Therapie entweder im „Richtlinien-Verfahren“ bewilligt worden war, - bei denen der Krankheitswert also per Gutachterverfahren bestätigt war, oder aber bei denen es sich um klinisch untergebrachte Langzeitpatienten handelte. Nach Vorliegen der ersten klinischen Daten zum Messzeitpunkt t1 konnten wir unsere Patienten „empirisch“ auf ihren Krankheitszustand hin überprüfen, indem wir die bei ihnen gemessenen Werte mit den Werten der Eichpopulationen der verwendeten Messinstrumente verglichen. Dabei ergab sich, dass sich unsere Patienten zu Beginn der Behandlung durchaus auf dem Krankheitsniveau der Eichstichprobe ‚kranker Patienten‘ befanden und sich nach dreijähriger Behandlungszeit den Mittelwerten der ‚gesunden Eichstichproben‘ annäherten (vgl. Kapitel 10.3.4.1 – 10.3.4.3).

### 5.5 Geographische Ausweitung der Untersuchung

Viele Untersuchungen zur Wirksamkeit von Psychotherapie sind kommunal gebunden, d.h., es wird eine geringe Anzahl von Patienten aus der direkten Umgebung des Forschers, der dann meist auch noch der Therapeut ist, in die Untersuchung hineingenommen. Wir wollten auch in dieser Frage von vorne herein methodologischen Zweifeln vorbeugen und entschlossen uns, bundesweit alle Therapeuten der DGIP anzu-

schreiben und um deren Mitwirkung zu bitten. Auch hofften wir, so eine möglichst gute Zufallsstichprobe zu erhalten, in der sich Ausbildungsunterschiede, kommunale Unterschiede, Abweichungen der diagnostischen Einschätzung u.a. gegenseitig aufheben.

#### 5.6 Ethische Einwände hinsichtlich der Bildung der Kontrollgruppe

Zum Standard der „Methodik der empirischen Sozialforschung“ gehört die Bildung eines quasi-experimentellen Untersuchungsdesigns unter Einbeziehung einer Kontrollgruppe. In den meisten Veröffentlichungen zur Wirksamkeit von Psychotherapie wird die Kontrollgruppe einerseits gefordert, andererseits verworfen. Dabei wird sehr gerne und geradezu monoton auf das ethische Problem hingewiesen, Patienten nicht warten lassen zu dürfen, weswegen dann die Forderung nach einer Kontrollgruppe in der Regel stillschweigend und mit dem Ausdruck größten Bedauerns fallen gelassen wird. Wir entschlossen uns, dies nicht zu tun, sondern trotz ethischer Bedenken der Fachwelt auf die Bildung der Kontrollgruppe nicht zu verzichten. Wir denken auch, dass wir dies ethisch vertreten können. Überall gibt es wartende Patienten. Aus Finanzierungsgründen oder anderen Begrenzungen müssen Patienten oft monatelang oder jahrelang auf nötige Operationen warten; Ärzte fahren in die Weihnachtsferien, obwohl Patienten dringend operiert werden müssten; die psychiatrischen Kliniken sind voll von schlecht oder psychotherapeutisch unversorgten Patienten, ohne dass sich dadurch irgend ein Therapeut bemüht sähe, aus ethischen Gründen abhelfend einzugreifen und sich über die Dienstzeit hinaus zu engagieren. Niedergelassene Psychotherapeuten weisen Patienten ebenfalls aus den verschiedensten Motiven, z.B. auch einfach aus Kapazitätsgründen ab, ohne sich ethische Gedanken darüber zu machen, ob und wie diese Patienten anderweitig versorgt werden; und es gibt psychotherapeutische Praxen, die führen mit Zustimmung der betroffenen Patienten Wartelisten. Warum sollten wir dann nicht diese in den Wartelisten geführten Patienten zu ihrer Befindlichkeit befragen dürfen. Das Argument, man dürfe aus ethischen Gründen keinen Patienten warten lassen, ist praktisch nicht durchzuhalten. Sicherlich haben wir die beteiligten Therapeuten gebeten, Patienten für die Kontrollgruppe zu gewinnen, jedoch war von vorne herein klar, dass kein Patient nur deshalb auf die Warteliste verschoben werden sollte, um ihn als Mitglied der Kontrollgruppe missbrauchen zu können. Die Patienten wurden auch nicht künstlich zu Studienzwecken in der Kontrollgruppe gehalten, sondern verließen diese Gruppe zu dem natürlichen Zeitpunkt ihres Therapiebeginns, - es handelte sich eben um eine „naturalistische“ Kontrollgruppe im Rahmen einer „naturalistischen“ Gesamtuntersuchung. Dass die Patienten von der KG in die UG gewechselt waren, erfuhr die Forschungsgruppe immer erst über einen rücklaufenden Fragebogen einer Wiederholungsmessung zum Zeitpunkt „tx“, in dem dann die Frage nach den zwischenzeitlich stattgefundenen Therapiesitzungen mit einer Anzahl „y“ beantwortet wurde.

#### 5.7 Langzeitstudie und Katamnese

Viele Studien leiden darunter, dass die untersuchte Population nur im Rahmen von Kurzzeittherapien oder eben nur über eine bestimmte Strecke einer Langzeittherapie untersucht wird. Wahrscheinlich steckt hier das Forscherinteresse dahinter, möglichst schnell zu ‚irgendwelchen‘ Ergebnissen zu kommen, Hauptsache ein Ergebnis, - und Hauptsache schnell. Die hierbei erzielten Forschungsergebnisse sagen dann in der Regel

nichts mehr über Langzeitpsychotherapien, sondern vielleicht noch etwas über Kurzzeitpsychotherapien aus. Demgegenüber hatten wir uns entschieden, eine wirkliche Langzeitstudie über eine Beobachtungsdauer von drei Jahren psychoanalytischer Langzeitpsychotherapien mit mehreren Messzeitpunkten und einer zusätzlichen katamnestischen Nachuntersuchung im zeitlichen Abstand von einem Jahr nach Abschluss der Therapien vorzulegen.

Wir wurden mehrmals von Kollegen gedrängt, schon einmal ‚erste Ergebnisse‘ unserer Untersuchung bekannt zu machen, haben dies aber aus methodologischen Gründen, d.h. zur Vermeidung eines Rückkopplungseffekts auf die beteiligten Therapeuten und die nicht kontrollierbaren Übertragungswirkungen auf die Patienten, also zur Vermeidung empiristisch-positivistisch-methodologischer Vorhaltungen hinsichtlich möglicher Verfälschungen des Untersuchungsergebnisses, strengstens unterlassen.

#### 5.8 Freiheit von den Finanzierungspartnern

Forschung kostet Geld, - diese Untersuchung kostete ca. 75.000,-- DM und hätte sicherlich noch weit mehr gekostet, wenn wir uns nicht immer wieder eingegrenzt hätten. Wir wollten dieses Geld natürlich so weit wie möglich nicht aus Eigenmitteln aufbringen. Wir wandten uns deshalb an die DGIP sowie an die angeschlossenen Ausbildungsinstitute. Diese finanzierten unsere Forschung, - eher zaghaft als großzügig -, aber sie gaben. Aber damit war natürlich auch ein Interesse verbunden. Die Geldgeber wollten schon sehr frühzeitig ‚erste Ergebnisse‘ mitgeteilt bekommen. Neben diesem reinen Wissens-Interesse mag wohl unterschwellig auch ein mehr oder weniger bewusstes oder unbewusstes Auftrags-Interesse, nämlich die Wirksamkeit der „Individualpsychologisch-Psychoanalytische Psychotherapie“ nachzuweisen, - und das noch möglichst schnell -, vorhanden gewesen sein. Für uns war mit diesen Erwartungen das Risiko verbunden, die Freiheit der Forschung zu verlieren. Wir legten großen Wert darauf, uns mit der Bewilligung von Geld nicht verpflichten zu lassen. Als schließlich die weitere Förderung vom Abschluss eines einklagbaren „Werkvertrags“ abhängig gemacht wurde, entschlossen wir uns zur privaten Weiterfinanzierung. Dies schien uns aus dem methodologischen Grundsatz der Sicherung der Unabhängigkeit als Grundvoraussetzung von (im empiristischen Sinne gemeinter) Objektivität erforderlich.

#### 5.9 Anonymität der Patienten

In vielen Wirksamkeitsuntersuchungen nehmen die Untersucher direkten Kontakt mit den Patienten auf, weil das Studiendesign ohne oder trotz Problembewusstsein so gewählt ist, dass dies nicht zu vermeiden ist. Häufig ist dies schon deshalb nötig, weil man sich im Studiendesign darauf festgelegt hat, qualitative Daten zu erheben. Hieran kann aus einer Argumentationssicht wie der „Methodik der empirischen Sozialforschung“ erhebliche methodologische Kritik angebracht werden. Das Entscheidende ist ja, dass diese qualitativen Befragungen die Anonymität des befragten Patienten aufheben und dann, - im positivistisch verstandenen Sinne -, die Gefahr besteht, dass der Patient sich aus der zwangsläufig entstehenden (wir würden sagen) ‚Übertragungsbeziehung zum Forscher‘ nicht mehr völlig frei und unabhängig in der ‚Abgabe seiner Daten‘ fühlt, sondern dass sich der sogenannte „Pygmalion-Effekt“ aufbaut. Der Kontakt des Forschers mit dem Patienten bringt aus empiristischer Sicht unkontrollierbare Störvariablen in das Forschungsgeschehen und ist

deshalb als methodologisch bedenklich abzulehnen. Diese methodischen Überlegungen hatten für uns die Konsequenz, weder die Therapeuten als ‚Forscher vor Ort‘ einzusetzen, noch selbst als Forscher direkten persönlichen Kontakt mit dem Patienten aufzunehmen, sondern das gesamte Forschungsverfahren auf anonyme Fragebogenaktionen zu begrenzen und das Forschungsdesign so aufzubauen, dass den Patienten die zugesagte Anonymität augenfällig gemacht wurde. Die Patienten erhielten die Anschreiben der Forschungsgruppe einschließlich aller Testunterlagen und Verhaltensanweisungen zwar von ihrem jeweiligen Therapeuten ausgehändigt, - damit war jedoch die Beteiligung der Therapeuten bereits beendet. Die Patienten hatten dann zwar die Möglichkeit, bei Bedarf über die im Briefkopf der Begleitschreiben genannte Kontaktadresse oder telefonisch mit der Forschungsgruppe Kontakt aufzunehmen (was jedoch nicht vorkam), diese Kontaktaufnahmemöglichkeit galt jedoch umgekehrt nicht für die Forschungsgruppe. Die Rückleitung der ausgefüllten Fragebögen an die Forschungsgruppe vollzog sich aus Gründen der Anonymität nicht über die Therapeuten, sondern auf postalischem Wege vom Patienten direkt an die Forschungsgruppe. Die Rücksendeumschläge waren von der Forschungsgruppe ausreichend vorfrankiert. Die Therapeuten waren in diesem Verfahren noch nicht einmal informiert, ob ihre Patienten die Fragebögen überhaupt ausfüllten und an die Forschungsgruppe zurück schickten. Um die Therapeuten aus dem Geschehen wirklich so weit wie möglich ‚herauszuhalten‘ und alle Verwicklungen für die Psychotherapien, aber auch methodologische Verstrickungen für die Untersuchung zu vermeiden, haben wir bei den Therapeuten auch nicht nachgefragt, wenn entweder einzelne Items oder sogar ganze Fragebogensendungen ihrer Patienten ausblieben. Wir vermieden so jedwede Verwicklung für die Therapeuten und jedweden „Druck“ auf die Patienten. Da aber alle Patienten zum nächsten Messzeitpunkt trotz ausgebliebener Rücksendungen erneut turnusmäßig über ihre Therapeuten angeschrieben wurden, ergab es sich, dass dann häufig zu den späteren Testzeitpunkten wieder Material bei uns einging, so dass die Datensätze häufig lückenhaft sind, was sich in den schwankenden „N“-Werten niederschlägt (vgl. Auswertungstabellen im Anhang).

#### 5.10 Handhabung der Rohwerte als objektive „Daten“

Der Verzicht auf ein Studiendesign, das eine Kontaktaufnahme zum Patienten erforderlich gemacht hätte, brachte es mit sich, alle ‚Kreuzchen‘ eines Patienten in einem Fragebogen unhinterfragt und im positivistischen Sinne als ‚Daten‘ zu werten. „Als Daten lassen sich nun beobachtbare Merkmalsausprägungen auf Merkmalsdimensionen von Untersuchungseinheiten bezeichnen; sie stellen nichts anderes dar, als einen jeweils bestimmten Wert auf einer Variablen.“ (Mayntz u.a., 1972, S. 35) „Zu Daten werden Beobachtungen ... also erst, wenn sie begrifflich so strukturiert sind, dass sie als registrierte Beobachtung eine Identifizierung und Klassifizierung der Untersuchungsobjekte nach ihren für die Beantwortung der Forschungsfrage relevanten Eigenschaftsdimension ermöglichen.“ (Mayntz u.a., 1972, S. 33) Daten sind also hier Ausprägungen auf einer empirischen Skala, nicht jedoch inhaltlich deutbare Sinngrößen. Die Diskussion über Fragen wie z.B., ob der Patient die Frage richtig oder falsch verstanden haben könnte, wie sein Antwortverhalten auf der Folie seiner Persönlichkeit gesehen werden kann oder ob sein Antwortverhalten in Hinblick auf unbewusste Motivationen heraus sogar als zweifelhaft angesehen werden muss, bleiben undiskutiert. Lediglich das nach au-



ßen gewandte und in diesem Sinne ‚positiv beobachtbare‘ Antwortverhalten ist relevant, - und wird durch die unhinterfragende Akzeptanz durch die Forscher zu einem Fakt, zu einem Datum.

#### 5.11 Stichprobenauswahl

Es stellt sich die Frage, ob die von uns untersuchte Patientengruppe Repräsentativitätskriterien im Sinne eines verkleinerten Abbildes der bundesdeutschen Patientenlandschaft genügt. Streng genommen wissen wir nicht, ob unsere Patientengruppe diesen Anforderungen genügt, und wir haben zur Erreichung dieses Zieles auch keine weiteren Anstrengungen unternommen. Dies deshalb, weil wir ja bereits große Schwierigkeiten hatten, die bundesweit angeschriebenen Therapeuten überhaupt zur Mitarbeit zu gewinnen und ein genügend großen „N“ zu erreichen. Wir konnten uns von daher nicht darum kümmern, ob die eingebrachten Patienten Repräsentativitätskriterien genügen oder nicht. Herausgekommen ist eine zufällige Untersuchungsgruppe und eine zufällige Kontrollgruppe, von denen wir auf Grund der Zufälligkeit ihres Zustandekommens vermuten, dass dieses Gesamtkollektiv ein hinreichend repräsentatives Abbild des bundesrepublikanischen „Patientenkollektivs“ darstellt, - belegen können wir dies nicht.

#### 5.12 Entscheidungsgründe für die angewandten Messverfahren

Hinsichtlich der Verwendung von Persönlichkeitsfragebögen weist Kubinger (1999) kritisch darauf hin, dass „selbst in Situationen, in denen die mit dem Testergebnis verbundenen Konsequenzen für die Testperson ziemlich gering, jedenfalls keineswegs existentiell sind, die Antworten in Richtung sozialer Erwünschtheit (bzw. in Richtung der Maximierung des individuellen Nutzens) gegeben“ (Kubinger, 1999, S. 135) werden. „Abgesehen davon, dass damit die Aussagekraft von Persönlichkeitsfragebögen äußerst gering ist, ‚verbietet‘ sich prinzipiell die ‚Idee‘, mit ihnen eindimensional messen zu wollen: Immer ist die eigentlich interessierende Eigenschaft von einer zweiten, nämlich der, ‚unwahr zu antworten‘, überlagert.“ (Kubinger, 1999, S. 135). Eine Ausnahme stellt nach Meinung Kubingers „vielleicht die Anwendung von Persönlichkeitsfragebögen im klinischen Bereich dar. Ist der Leidensdruck der Testperson genügend groß und erwartet sie sich gerade durch die psychologische Untersuchung Erkenntnisse, die ihr (direkt oder indirekt) zu Hilfe kommen, so ist eine authentische Beantwortung wahrscheinlich.“ (Kubinger, 1999, S. 135) Allerdings müsse „befürchtet werden, dass es selbst hier zu absichtlichen Täuschungen kommt.“ (Kubinger, 1999, S. 135)

Wir entschlossen uns, als Verfahren zur Messung der Patientenbefindlichkeit den Gießener Beschwerdebogen (GBB), den Minnesota Multiphasic Personality Inventory (Deutsche) Kurzform (MMPI/K) und den Veränderungsfragebogen des Erlebens und Verhaltens (VEV) einzusetzen und die Unterschiede der Messergebnisse zwischen den verschiedenen Messzeitpunkten als ein Maß für die Therapieeffektivität anzusehen. Die Gründe hierfür waren wie folgt:

##### 5.12.1 Fragen zur Operationalisierung, Objektivität, Validität und Reliabilität

Wir wollen „Therapieerfolg“ messen. Stellt sich die Frage nach der Operationalisierung von „Therapieerfolg“. Operationalisierung meint die „Festlegung von Forschungsoperationen, die das Vorliegen eines begrifflich bezeichneten Phänomens entscheidbar machen sollen.“ (Mayntz u.a., 1972, S. 18) Die Notwendigkeit der

Operationalisierung hat damit zu tun, dass das Phänomen „Therapieerfolg“ nicht unmittelbar wahrgenommen werden kann. Der Umstand, dass wir uns zu einer empirischen Studie gemäß den Regeln der empirischen Sozialforschung entschieden haben, beinhaltet Implikationen und Konsequenzen für die Operationalisierung von „Therapieerfolg“, nämlich dass wir uns in Hinblick auf den Einsatz objektiver Messverfahren nur noch „auf der Ebene der Krankheitssymptomatik“ (Langenmayr und Kosfelder, 1998, S. 98) bewegen können. Bezogen auf diese „Ebene des Therapieerfolgs“ müssen wir nun angeben, wie wir den „Erfolg“ definieren. Dabei bieten sich nach Langenmayr und Kosfelder (1998) zwei Möglichkeiten an, nämlich zum einen die gemessene Veränderung zwischen Therapiebeginn und Therapieende und zum anderen das Erreichen eines normativen Ziels wie z.B. die Unterschreitung klinisch relevanter Messwerte in den herangezogenen Messverfahren. In einem weiteren Schritt müssen diese Messverfahren festgelegt werden, womit die Operationalisierung des „Therapieerfolgs“ abgeschlossen ist.

An dieser Stelle stellt sich die Frage, nach welchen Kriterien die Messverfahren ausgewählt werden. Allgemein kann gesagt werden, die Messinstrumente sollen adäquat sein, Therapieerfolg überhaupt messen zu können. Adäquat ist ein Messinstrument nach messtheoretischer Lesart dann, wenn es in der Lage ist, die zu beobachtende Veränderung auf einer Skala durch unterschiedliche Messwerte abzubilden. Um hierzu in der Lage zu sein, müssen die vorgesehenen Messinstrumente bestimmten messtheoretischen Gütekriterien genügen. Hierzu zählt man gemeinhin erstens die Objektivität, d.h., das Messinstrument muss die Möglichkeit einer subjektiven Verfälschung des Messergebnisses durch die Person des Messenden ausschließen. Zweitens muss das Messinstrument das messen, was es zu messen vorgibt bzw. was es messen soll (Validität). Drittens muss das Messinstrument selbst zuverlässig, d.h. ‚in sich selbst so stabil und veränderungsresistent‘ sein, dass der Messvorgang prinzipiell jederzeit bei unterstellter Merkmalskonstanz mit gleichem Messergebnis wiederholbar ist (Reliabilität).

Was unsere Untersuchung anbelangt, so stellt sich an dieser Stelle die Frage nach der Objektivität, Tauglichkeit/Validität und Messgenauigkeit/Reliabilität der ausgewählten Messverfahren GBB, MMPI(K), VEV, Helping Alliance Questionnaire (HAQ) und der Intrex-Kurzform. Die beiden letztgenannten Verfahren beinhalten Einschätzungsskalen und sind für den katamnestischen Teil der Untersuchung vorgesehen. Die drei erstgenannten Verfahren kamen jedoch im eigentlichen Kernbereich der Studie zum Einsatz. Die Frage der Eignung der ausgewählten Verfahren hinsichtlich der Erfassung des Untersuchungsgegenstandes „Therapieerfolg“ soll sich deshalb an dieser Stelle auf diese drei Verfahren beschränken. Die Darstellung des HAQ sowie der Intrex-Kurzform erfolgt im katamnestischen Teil dieser Studie an gesonderter Stelle (vgl. unten).

Während die Objektivität des GBB, MMPI(K) und des VEV unstrittig ist, stellen sich Validität und Reliabilität etwas problematischer dar. Uns wurde nämlich bereits im Vorfeld unserer Untersuchung zweifelnd vorgehalten, die genannten Verfahren seien ursprünglich ja gar nicht zur Messung von Psychotherapieeffektivität konzipiert worden, weswegen ihr Einsatz eher kritisch zu sehen sei. Es sei daher sinnvoller, andere Messverfahren zum Einsatz zu bringen, die für den Messzweck ‚Therapieerfolg‘ entwickelt worden seien. Hiermit ist vor allem die Frage nach der Validität (Tauglichkeit) der von uns eingesetzten Verfahren angesprochen. Es geht um die Frage, ob wir die „Wirksamkeit von Psychotherapie“ nur mittels zu diesem Zweck konstruierter

Verfahren sinnvoll messen können, oder ob auch ursprünglich zu anderen Messzwecken konstruierte Verfahren hier tauglich sind.

a) Validität: Die grundsätzliche Frage lautet: Sind die von uns verwendeten Verfahren überhaupt vom Grundsatz her geeignet/tauglich, Wirksamkeit von Psychotherapie zu messen, da sie doch ursprünglich gar nicht zu diesem Zweck konstruiert wurden. Vielmehr messen diese Verfahren Ausprägungen auf bestimmten klinischen Skalen. Messen wir also überhaupt „Wirkung von Psychotherapie“ oder messen wir etwas anderes? Zunächst einmal messen wir Ausprägungen auf verschiedenen klinischen Skalen. Allerdings wird von uns die „Wirkung von Psychotherapie“ als die Veränderung auf den klinischen Skalen zwischen verschiedenen Messzeitpunkten definiert. Dennoch bleibt die Grundfrage bestehen, ob wir mittels dieser operationalen Definition auch wirklich die Wirkung von Psychotherapie erfassen können oder nicht. Bei vordergründigem Hinsehen scheint sich hierbei zunächst ein altes Grundsatzproblem entgegenzutreten, das in der Geschichte der Psychologie wohl erstmals an den „Intelligenztests“ durchdiskutiert wurde. Die Frage lautete, ob der Intelligenztest überhaupt mißt, was er zu messen vorgibt, nämlich Intelligenz. Da es keine objektive und empirische Einigkeit darüber geben kann, was Intelligenz ist, blieb dieser Streit fruchtlos und führte schließlich zu der positivistischen Extremposition: ‚Intelligenz ist das, was der Intelligenztest misst‘. Zunächst scheint es nun so, als würden wir mit der Frage nach der „Wirksamkeit von Psychotherapie“ in eine ähnliche Problemlage hineingezogen: messen die ausgewählten Messverfahren überhaupt Therapieeffektivität, - sind sie tauglich (valide), Veränderungen zu messen?

Da die Verfahren ja nicht vorgeben, Therapieeffektivität zu messen und zu diesem Zweck ja ursprünglich auch nicht konzipiert sind, muss die Frage zunächst einmal verneint werden. Die Frage der Validität stellt sich in unserer Untersuchung jedoch ganz anders. Wir haben nämlich gar nicht zu klären, welches Verfahren ‚Therapieerfolg‘ misst, und zwar deshalb nicht, weil es den ‚Therapieerfolg‘ als eine positive Eigenschaft einer Person ja gar nicht gibt. Uns interessiert doch nicht die einmalige Messung einer Eigenschaft, wie z.B. eine einmalige Messung der Intelligenz, sondern uns interessiert die Veränderung von Ausprägungen zwischen zwei Messzeitpunkten, - nämlich ob es zwischen zwei zeitlich versetzten Messungen, zwischen denen so etwas stattgefunden hat, was wir Therapie nennen, Unterschiede in den Ausprägungen von klinisch relevanten Störungswerten gibt.

Bezogen noch einmal auf die „Intelligenzthematik“ würde uns hier ja nur interessieren, ob zwischen zwei zeitlich versetzten Messungen, zwischen denen so etwas stattgefunden hat, was wir z.B. „Beschulung“ nennen, Unterschiede in den Messwerten eines Intelligenztests gibt, egal ob der Intelligenztest nun wirklich Intelligenz misst oder nicht und egal was er überhaupt misst. Hätten wir also die Frage zu klären, ob Psychotherapie die Intelligenz verändert, müssten wir uns nicht auf die Suche nach einem neuen Test für die ‚Messung der Therapiewirkung auf die Intelligenz‘ machen, sondern wir würden einen bereits existierenden Intelligenztest heranziehen, der hinsichtlich der Frage der Intelligenzmessung gute Validitätswerte aufzuweisen hätte und deshalb gemeinhin anerkannt wäre, - und wir würden diesen Intelligenztest vor und nach der Psychotherapie einsetzen und den Messunterschied mit der zwischenzeitlich stattgefundenen Psychotherapie in einen Zusammenhang bringen. Uns würde also bei dem Intelligenztest zunächst nur die Frage nach der Validität

hinsichtlich der Intelligenz, nicht aber hinsichtlich des Messzwecks „Therapieerfolg“ interessieren. Der Messzweck „Therapieerfolg“ ließe sich dann sehr leicht über die Feststellung der Unterschiedlichkeit der Messwerte vor und nach Einführung der Variablen ‚Psychotherapie‘ erfassen. So gesehen ist für uns die Frage, ob die von uns angewandten Verfahren überhaupt „Therapieerfolg“ messen, also die Frage nach der Validität hinsichtlich des „Therapieerfolgs“ gar nicht von Bedeutung. Die Frage der Validität hinsichtlich der Messung des „Therapieerfolgs“ entscheidet sich an ganz anderer Stelle, nämlich wenn es darum geht, alle anderen Variablen so zu kontrollieren, dass die Messdifferenz einer Prä-Post-Messung auch wirklich nur auf Psychotherapie und nicht auf andere unkontrolliert wirkende Variablen zurückgeführt werden kann. Die von uns eingesetzten Verfahren müssen also gar nicht „Therapieerfolg“ messen, - müssen diesbezüglich also auch gar nicht valide sein. Dennoch ist die Validität der eingesetzten Verfahren nicht bedeutungslos: sie sollen Ausprägungen auf klinischen relevanten Skalen möglichst valide messen, damit der behauptete Zusammenhang zwischen Psychotherapie und klinischen Bildern überhaupt fassbar wird und hierüber Aussagen möglich werden. Die Validität des Messinstruments interessiert uns sogar nur und ausschließlich in Hinblick auf die klinischen Skalen, die durch Psychotherapie verändert werden sollen, - denn es interessiert uns ja nicht „irgendeine beliebige“ Veränderung durch Psychotherapie, sondern nur jene Veränderung, die sich in klinisch relevanten Bereichen abspielt. Uns interessiert z.B. nicht die Veränderung des Lebensalters zwischen zwei Messzeitpunkten, zwischen denen Psychotherapie stattgefunden hat. Zwar misst die zu zwei unterschiedlichen Messzeitpunkten der Psychotherapie vorgelegte Frage „Wie alt sind Sie?“ sehr valide die Veränderung des Alters, doch würde uns diese sehr hohe „Veränderungsvalidität“ ja gar nicht interessieren, weil sie erstens nicht durch Psychotherapie erzielt wurde und sich zweitens auch nicht auf einen klinisch relevanten Sachverhalt bezieht. Die Bedeutung der Validität des Messinstruments kann uns deshalb lediglich in Hinblick auf die klinischen Skalen, die durch Psychotherapie verändert werden sollen, interessieren.

Damit aber sind wir auf der Ebene nicht mehr der Prüfung der Validität unserer Messverfahren hinsichtlich des „Therapieerfolgs“, sondern auf der Ebene der Prüfung der Validität der Messinstrumente hinsichtlich ihres je eigenen Messzwecks, die uns für unsere Studie völlig ausreicht.

b) Reliabilität: Wichtiger als die Validität ist für unsere Untersuchung, dass wir erstens zu den verschiedenen Messzeitpunkten immer wieder das gleiche Messinstrument verwenden, und dass wir uns zweitens darauf verlassen können, dass dieses Messinstrument zwischen den beiden Messzeitpunkten konstant und unverändert geblieben ist. Denn nur unter der Bedingung der Konstanz des Messinstruments sind Messungen ja überhaupt nur miteinander vergleichbar und nur so ist ein etwaiger Messunterschied auf eine Veränderung am gemessenen Objekt und nicht auf eine Veränderung am Messinstrument zurückzuführen. Dies aber ist die Frage nach der Reliabilität: Nur wenn das Messinstrument sich nicht verändert, kann eine Messung reliabel, also wiederholbar, in dem Sinne sein, dass eine gleichgebliebene oder veränderte Messwertausprägung dem gemessenen Gegenstand und nicht dem Messinstrument zugeschrieben werden muss. Da unsere Absicht darin besteht, Veränderungen zu messen, ist für uns von zentraler Bedeutung, ob unsere Messverfahren zwischen den gewählten Messzeitpunkten konstant bleiben, ob wir also bei wiederholten Messungen am, - hier einmal unterstellt -, „gleich gebliebenen Objekt“ das gleiche Messergebnis feststellen können. Dann

kann gelten: Therapiewirkung ist der Unterschied der Messausprägungen/Messwerte, den unsere verwendeten Messverfahren zu verschiedenen Zeitpunkten messen, vorausgesetzt, dass ansonsten alle anderen Variablen konstant gehalten werden konnten und nur die ‚Therapie-Variable‘ variiert.

5.12.2 Der „Giessener Beschwerdebogen“ (GBB) wird im klinischen Bereich immer wieder eingesetzt, ist also der Fachwelt bekannt, und ist im übrigen auch bereits bei Therapieeffektivitätsstudien eingesetzt worden (vgl. unten). Darüber hinaus hat der GBB eine hohe ‚face-validity‘ für die Patienten, - soll heißen, dass die große Nähe zu den üblichen Beschwerden, welche die Patienten zunächst zu ihrem Hausarzt und dann in psychotherapeutische Praxen führen, den Patienten das Gefühl geben, dass die Untersuchung wirklich eine Relevanz hinsichtlich ihrer Zustandsbilder hat, - was sicherlich ein eher günstiger Ausgangspunkt für die Erzeugung einer ernsthaften und mitwirkungsbereiten Verfassung hinsichtlich unserer Untersuchung ist.

Für den GBB spricht weiterhin, dass er somatoforme Beschwerden sowohl einzeln als auch in Beschwerdekomplexen erfassen kann. Da somatoforme Erkrankungen im Verlauf zumeist polysymptomatisch fluktuieren, sollte die Fokussierung auf ein einzelnes Symptom nur mit einer gewissen Vorsicht erfolgen. Rudolf (1998) empfiehlt deshalb, aggregierende diagnostische Verfahren einzusetzen. Dem wird im GBB vor allem mit der zusammenfassenden Skala Nr. 5 (vgl. unten) entsprochen, - eine Skala, die deshalb in unseren Auswertungen eine bevorzugte Rolle spielen soll.

Der Gießener Beschwerdebogen (GBB) besteht aus fünf Skalen:

Skala 1: Erschöpfung, bestehend aus Items, die auf Schwächegefühl, Schlafbedürfnis, Erschöpfbarkeit, Müdigkeit, Benommenheit und Mattigkeit abzielen,

Skala 2: Magenbeschwerden, bestehend aus Items, die Völlegefühl, Erbrechen, Übelkeit, Aufstoßen, Sodbrennen und Magenschmerzen abfragen,

Skala 3: Gliederschmerzen, inhaltlich durch Items hinsichtlich Gliederschmerzen, Rückenschmerzen, Nackenschmerzen, Kopfschmerzen, Müdigkeit in den Beinen und Druckgefühl im Kopf dargestellt,

Skala 4: Herzbeschwerden, welche die Symptome Herzklopfen, Schwindelgefühl, Kloßgefühl in der Brust, Stiche in der Brust, Atemnot und Herzbeschwerden umfasst, sowie die zusammenfassende

Gesamtskala 5: Beschwerdedruck, welche die Items aller vorgenannten Skalen umschließt (vgl. Brähler und Scheer, 1995, S. 23).

Obwohl die Testautoren Brähler und Scheer die psychische Bedingtheit der körperlichen Beschwerden sehen und einen klar psychosomatischen Ansatz verfolgen, haben sie bei der Konstruktion des GBB die direkte Ansprache der ‚psychischen‘ Beschwerden weggelassen und verweisen darauf, dass „für diesen Bereich durch psychologisch-psychiatrische Tests, wie den MMPI, breiter gefächerte Diagnoseinstrumente zur Verfügung stehen“. (Brähler und Scheer, 1995, S. 12) Durch den Bezug auf körperliche/psychosomatische Beschwerden erhält der GBB die von uns für unsere Untersuchung so geschätzte ‚face-validity‘, sprich ‚Patientennähe‘.

Gütekriterien: Brickenkamp bescheinigt dem GBB „vollständige Durchführungs- und Auswertungsobjektivität“ (Brickenkamp, 1997, S. 756). Was den Test-Retest-Reliabilitätswert anbelangt, so referiert Brickenkamp für alle fünf Skalen mittlere Korrelationen von 0,43 bis 0,71, für die zusammenfassende Skala „Beschwerdedruck“ Werte von 0,61 bis 0,93. Bei der Split-Half-Methode seien für die ersten vier Skalen Werte von 0,61 bis 0,73 und für die zusammenfassende Gesamtskala „Beschwerdedruck“ Werte von 0,85 und 0,84 gemessen worden. Die Validität des GBB wurde mittels Vergleichen zu Patienten mit fest umschriebenen klinischen Beschwerdebildern sowie mit Extremgruppenvergleichen geprüft, wobei von Brickenkamp (1997) hochsignifikante Ergebnisse im Sinne der Validitätserwartung referiert werden.

Die Testautoren haben sehr ausführliche Erhebungen von Normwerten durchgeführt (Brähler u.a., 1995, S. 27 ff.). Sie wurden anhand einer repräsentativen Eichstichprobe der Bevölkerung der Bundesrepublik Deutschland sowie anhand einer Patientenstichprobe erstellt. Diese Normwerte sind für unsere Untersuchung insofern wichtig, als wir darüber Vergleiche zu unserer Untersuchungs- und Kontrollgruppe anstellen und so Aussagen über den Grad der ‚Gestörtheit‘ der von uns untersuchten Patienten gewinnen können.

Referenzuntersuchungen, in denen der GBB zum Einsatz kam: a) Zielke, M.: Wirksamkeit stationärer Verhaltenstherapie (1993), b) Wunderlich, G.: Verringerung von Körperbeschwerden durch Psychotherapie (1993), c) Egle U., Heucher, K., Hoffmann, S., Porsch, U.: Psychoanalytisch orientierte Gruppentherapie mit psychogenen Schmerzpatienten (1992), d) Kordy, H., von Rad, M., Senf, W.: Therapeutische Faktoren bei stationärer Psychotherapie (1990), e) Schulz, H., Lang, K., Lotz-Rambaldi, W., Bürger, W., Koch, U.: Analyse von Behandlungsabbrüchen in der stationären psychosomatischen Rehabilitation anhand von Basisdokumentationen zweier Klinikträger (1999), f) Broda, M., Lamberti, J., Dinger-Broda, A., Klinkenberg, N.: Entwicklung eines Fragebogens zur Standardkatamnese für die stationäre Psychosomatische Rehabilitation (1999).

5.12.3 Das „Minnesota Multiphasic Personality Inventory“ (MMPI) sowie der MMPI/K (Kurzform) ist in den letzten Jahren zunehmend umstritten und es gibt in der Literatur eine Reihe von sich widersprechenden Aussagen zur Validität und Reliabilität dieser Verfahren. Wir haben uns dennoch für den Einsatz des MMPI(K) entschieden, - und zwar weil er eines der im deutschsprachigen Raum bekanntesten und am meisten untersuchten und eingesetzten klinischen Mess- und Testverfahren ist.

Der MMPI(K) besteht aus 221 Items, von denen Brickenkamp sagt, dass „aufgrund empirischer Untersuchungen ausgesagt werden kann, dass sie von den insgesamt 566 Items der Langfassung des MMPI den höchsten Beitrag zum Zustandekommen des MMPI-Profiles leisten“. (Brickenkamp, 1997, S.782) Die Items bilden zehn Persönlichkeitsbereiche ab, die gleichzeitig auch klinische Skalen darstellen, insofern nämlich der Höhe der erzielten Ausprägungen auf diesen Skalen klinische Zustandsbilder zugeordnet werden können. Im einzelnen handelt es sich hierbei um Hypochondrie, im Testhandbuch als „Hypochondriasis“ (HS) (Gehring u.a., 1993, S. 7) bezeichnet, um Depression (D), Hysterie (HY), Psychopathie (PP), Maskulin-Feminine Interessenskala (MF), Paranoia (PA), Psychasthenie (PT), Schizoidie (SC), Hypomanie (MA) und Soziale Introversion (SI). Darüber hinaus gibt es drei zusätzliche „Validitätsskalen“ (Brickenkamp, 1997, S. 782), die Aussagen über die Güte der zehn klinischen Skalen geben: Die Lie-Skala (L) zeigt den Grad der Haltung der Testperson an, sich in ein günstiges Licht zu setzen (Abwehrhaltung), die Correctionr-Skala (K)

gibt Auskunft über die Einstellung der Testperson zum Gesamttest und sagt so etwas über die Diskriminierungsschärfe der gemessenen Werte auf den klinischen Skalen aus, der F-Wert gibt Auskunft über die Frage, ob der Test kognitiv überhaupt richtig verstanden bzw. ausgefüllt wurde, - die Ergebnisse also überhaupt interpretierbar sind.

Gemäß Hathaway und McKinley finden sich bei stark neurotischen Patienten erhöhte Werte vor allem in den Skalen Hypochondriasis (HS), Depression (D) und Hysterie (HY), weshalb diese drei Skalen auch als „neurotische Trias“ (Hathaway u.a. 1991, S. 48) bezeichnet werden. Weiter wird ausgeführt, dass sich als Neuroseindikator meist auch erhöhter Psychastheniewert (PT) findet. Wir haben uns entschlossen, diese in der Literatur vorgefundene Skalengewichtung aufzugreifen und die drei Skalen der „neurotischen Trias“ in Verbindung mit dem Psychastheniewert (PT) in den Mittelpunkt auch unserer MMPI(K)-Auswertung zu stellen.

Gütekriterien: Brickenkamp hat keine Zweifel an der Objektivität des Tests hinsichtlich Durchführung und Auswertung. Weiterhin berichtet er von Reliabilitätskoeffizienten zwischen 0,861 und 0,968 bei der Retest-Methode und von 0,839 bis 0,974 nach der Split-Half-Methode (Brickenkamp, 1997, S. 783). Nach Brickenkamp liegt die Korrelation der Profile zwischen der Gesamtform und der Kurzform bei  $r = 0,803 \pm 0,154$  (S. 783). Die Auskünfte zur Validität sind gemäß Brickenkamp (1997) nicht ganz zufriedenstellend.

Normen: Das Handbuch zur Deutschen Kurzform werden Mittelwerte und Streuungen für die Normalpopulation (Gehring u.a., 1993, S. 64) sowie Mittelwertprofile für mehrere klinische Bilder angegeben (1993, S. 41 ff.), die als Vergleichswerte für die Einschätzung des ‚Gestörtheitsgrades‘ der von uns untersuchten Patienten herangezogen werden können.

Referenzuntersuchungen, in denen der MMPI zum Einsatz kam: a) Burbiel, I., Apfelthaler, R., Sandermann, G.: Stationäre Psychotherapie von Borderline-Patienten: Eine Prä/Post- und katamnestische Studie (1994), b) Burbiel, I., Apfelthaler, R., Schanne, U., Wolfrum, G.: Stationäre Psychotherapie der Psychosen; Eine testpsychologische katamnestische Untersuchung (1992), c) Burbiel, I., Apfelthaler, R., Schanne, U., Seidler, G., Wolfrum, G.: Die Effizienz stationärer Psychosentherapie in der Dynamisch-Psychiatrischen Klinik Menterschwaige München (1990), d) Paar, G., Kriebel, R., Noack, R.: Eine empirische Analyse zum Therapieabbruch bei stationärer Psychotherapie (1994), e) Strauss, B., Hess, H.: Interpersonale Probleme, interpersonale Orientierung und Behandlungserfolg nach stationärer Gruppenpsychotherapie (1993), f) Remschmidt, H., Wienand, F., Wewetzer, C.: Der Langzeitverlauf der Anorexia nervosa (1990).

5.12.4 Der „Veränderungsfragebogen des Erlebens und Verhaltens“ (VEV) ist deshalb mit in die Untersuchung hinein genommen worden, weil er von Grawe im „Forschungsgutachten“ als das Verfahren der Wahl zur direkten Messung von Therapieeffektivität genannt wird. Dies war uns Grund genug, dieses Verfahren auf jeden Fall in unserer Untersuchung zu verwenden. Es sollte nicht heißen, wir hätten uns gedrückt, das von Grawe bevorzugt ausgewiesene Messinstrument übergangen zu haben.

Der VEV erlaubt den Testpersonen, sich rückschauend hinsichtlich einer bei sich selbst wahrgenommenen Veränderung im Erleben und Verhalten quantitativ einzuschätzen. Die eingeschätzte Veränderung wird hinsichtlich der allgemeinen Erlebensweise, hinsichtlich sozialer Situationen und hinsichtlich ihrer Auswirkung in Leistungssituationen erfragt und auf einem Kontinuum der „Veränderung des Erlebens und Verhaltens“ mit

den Polen „Entspannung, Gelassenheit, Optimismus“ bzw. „Spannung, Unsicherheit, Pessimismus“ abgebildet. Die Stärke der selbsteingeschätzten Veränderung wird bei jedem Item mittels einer Skala von sieben Punkten ausgedrückt, die Punkte aller 42 Items werden zu einem Gesamtwert zusammengefasst.

Gütekriterien: Brickenkamp bescheinigt dem Verfahren Objektivität und eine „extrem hohe Messgenauigkeit“ (Brickenkamp, 1997, S. 811) bei Reliabilitätswerten zwischen  $r = 0,946$  und  $0,989$ . Kriterienbezogene Validitätskoeffizienten seien zwischen  $r = -0,29$  und  $r = 0,63$  und diskriminante Validitätskoeffizienten zwischen  $r = -0,11$  und  $r = -0,18$  ermittelt worden und auch faktorielle Validitätskoeffizienten seien befriedigend belegt. (Brickenkamp, 1997, S. 811)

Normen (referiert nach Brickenkamp, 1997, S. 812): Es liegen Normwerte für eine Therapie- und eine nicht behandelte Kontrollgruppe vor. Nichtveränderung wird angenommen, wenn alle Items mit einer mittleren Punktzahl versehen werden, was einem Gesamtpunktwert von 168 Punkten entspricht. Von positiven Veränderungen wird bei einem Gesamtwert von 187 bis 294 gesprochen, von negativen Veränderungen bei einem Gesamtpunktwert zwischen 149 und 42.

Referenzuntersuchungen, in denen der VEV zum Einsatz kam: a) Zielke, M.: Wirksamkeit stationärer Verhaltenstherapie (1973), b) Brockmann, J., Schlüter, T., Eckert, J.: Die Frankfurt-Hamburg-Langzeit-Psychotherapiestudie - Ergebnisse der Untersuchung psychoanalytisch orientierter und verhaltenstherapeutischer Langzeit-Psychotherapien in der Praxis niedergelassener Psychotherapeuten (2001), c) Kern, C., Küchenhoff, J.: Zur Evaluation der Psychotherapie im teilstationären Setting (2000).

### 5.13 Auswertungsvorentscheidungen

#### 5.13.1 Festlegung der auszuwertenden Messzeitpunkte

Vor Beginn der Studie wurden zunächst fünf Messzeitpunkte  $t_1$ ,  $t_2$ ,  $t_3$ ,  $t_4$  und  $t_5$  sowie eine katamnestiche Nachbefragung für den Zeitpunkt ‚ein Jahr nach Beendigung der jeweiligen Therapie‘ festgelegt. Die Erhebung der Daten wurde zu diesen Messzeitpunkten auch tatsächlich durchgeführt. Aus Kostengründen wurden dann jedoch nur die Daten der Messzeitpunkte  $t_1$ ,  $t_3$ ,  $t_5$  und die Katamnesedaten ausgewertet.

Was den VEV betrifft, so kam dieser zum Messzeitpunkt  $t_1$  noch nicht zum Einsatz. Dies deshalb, weil es sich beim VEV um einen Fragebogen handelt, bei dem der aktuelle Zustand zum Zeitpunkt der Befragung durch den Patienten mit einem früheren Zeitpunkt mittels einer subjektiven Veränderungseinschätzung verglichen wird. Weil zum Zeitpunkt der Erstbefragung aber naturgemäß noch kein rückblickender Bewertungszeitraum vorliegt, der eingeschätzt hätte werden können, kam der VEV zum ersten Mal zum Messzeitpunkt  $t_2$  zum Einsatz. Zur Auswertung kamen jedoch, wie oben bereits gesagt, aus Kostengründen lediglich die Daten von Messzeitpunkt  $t_3$  und  $t_5$ .

#### 5.13.2 Gründe für die Auswertung mittels t-Test und Varianzanalyse

Das t-Test-Verfahren „dient bei abhängigen Stichproben dem Mittelwertvergleich zweier Verteilungen, wobei zweidimensional normalverteilte metrische Daten vorausgesetzt werden. Es soll die Frage beantworten, ob sich die Mittelwerte der normalverteilten Grundgesamtheiten, aus denen die Stichproben (x) und (y) stammen, unterscheiden.“ (Clauß u.a., 1999, S. 235) Eine stichprobenartige Prüfung unserer Daten bestätigte



unsere Annahme der Normalverteilung unserer Daten, weshalb der t-Test als statistisches Prüfverfahren für den Mittelwertvergleich unserer Daten geeignet erschien. Wir entschieden uns weiter, mittels der t-Tests Signifikanzuntersuchungen der Mittelwertunterschiede im Vergleich Untersuchungs- und Kontrollgruppe (Querschnittuntersuchungen), aber auch innerhalb der jeweiligen Gruppe zwischen den aufeinander folgenden Messzeitpunkten (Längsschnittuntersuchungen) durchzuführen. Die so ermittelten Ergebnisse können sowohl etwas über den Therapieerfolg der Gesamtuntersuchungsgruppe im Vergleich zur Gesamtkontrollgruppe, als auch etwas über die zeitliche Entwicklung eines etwaigen Gesamttherapieerfolges in der Untersuchungsgruppe aussagen.

Ein Nachteil dieser t-Tests besteht darin, dass immer nur zwei Datengruppen miteinander verglichen werden, also nur Paarvergleiche gerechnet werden können. Zu einer integrierenden Verarbeitung unserer Daten im Sinne einer kombinierten Querschnitt-Längsschnitt-Untersuchung, sozusagen zur Verrechnung der Datengruppen ‚in einem Schritt‘ sind die t-Tests nicht in der Lage. Die Summation der einzeln durchgeführten t-Tests ergibt zwar ein Gesamtbild, doch bleibt dieses Gesamtbild notwendigerweise deskriptiv und damit auch spekulativ, - kommt also nicht zu einer mathematischen Verrechnung der einzelnen t-Tests. Um zu einer solchen mathematischen Verrechnung zu kommen, führen wir über die t-Tests hinausgehend mit unseren Datengruppen eine Varianzanalyse durch: Die Varianzanalyse erlaubt eine integrierte Fehlertermbestimmung von Querschnitt- und Längsschnittvergleich unserer Datengruppen (UG zu t1 und T3, KG zu t1 und t3).

Weiterhin interessierte uns die Frage, ob es signifikante Zusammenhänge zwischen einem etwaigen Therapieerfolg und bestimmten unterschiedlichen unabhängigen Variablen der Untersuchungsgruppe gibt. Zu diesen Variablen gehören die Diagnosegruppe der untersuchten Patienten, das Patientengeschlecht, das Patientenalter, die Sitzungsfrequenz, das Therapeutengeschlecht, das Therapeut-Patient-Geschlechter-Matching, der Grundberuf des Therapeuten (Arzt bzw. Psychologe), der Schulabschluss der Patienten, die Höhe des Anfangswiderstandes (operationalisiert im Anfangs-Lie des MMPI(K)), die Höhe der Anfangsdepression und die Höhe des anfänglichen Beschwerdedrucks. Zur Untersuchung des Einflusses dieser Variablen auf den Therapieerfolg planten wir, die Patienten der Untersuchungsgruppe hinsichtlich der genannten Variablen jeweils in sich voneinander unterscheidende Subgruppen aufzuteilen. Die so gebildeten Subgruppen sollten mit Hilfe von Signifikanztests hinsichtlich ihrer unterschiedlichen Entwicklung im Zeitverlauf miteinander verglichen werden. Auch zur Prüfung dieser Signifikanzen bedienten wir uns der Varianzanalyse. Die Begründung der Entscheidung zum Einsatz der Varianzanalyse liefern Clauß u.a. (1999, S. 313):

„Die Varianzanalyse ist ein statistisches Verfahren zur Untersuchung der Wirkung von Faktoren auf Versuchsergebnisse, das auf der Grundlage des geeigneten Vergleichens von Streuungen beruht. ... In Abhängigkeit von der Anzahl der Einflussgrößen, den sogenannten Faktoren, unterscheidet man die einfache, zweifache usw. Varianzanalyse.

Man unterscheidet nun das sogenannte Modell I, das einem globalen Mittelwertvergleich entspricht und von sogenannten festen Effekten ausgeht, und das Modell II, das die von den Faktoren erzeugte Variabilität untersucht und dabei zufällige Effekte zugrundelegt. ...

In der Psychologie wird die Varianzanalyse im wesentlichen zur Prüfung von Mittelwertunterschieden eingesetzt. Es werden dabei metrische Daten mit Normalverteilung sowie eine homogene Varianz der Stichprobenvariablen vorausgesetzt.“

In unserem Falle handelt es sich um die Untersuchung der Wirkung des Faktors „Therapieeinsatz“ auf das Versuchsergebnis einer Prä-Post-Anordnung zweier Stichproben mit unterstellter Normalverteilung und unterstellter homogener Varianz der Stichprobenvariablen.

Bei Vorliegen dieser Bedingungen schlagen auch Vorberg und Blankenberger (1999) die Varianzanalyse als geeignetes statistisches Prüfverfahren vor.

## **6. Studiendesign**

Unsere Untersuchung war von Anfang an mit Untersuchungsgruppe (UG) und Kontrollgruppe geplant. Ziel war es, in der UG mit 200 Patienten und in der KG mit 100 wartenden Patienten zu beginnen. Leider gelang es uns nicht, dieses Ziel zu erreichen. Obwohl immer noch Patienten nachgemeldet werden konnten, erreichte unsere Untersuchungsgruppe zum Messzeitpunkt t1 lediglich eine Größe von 144 Patienten und die Kontrollgruppe von 51 Patienten. Die Wartezeit für die Patienten sollte im Idealfall ein Jahr betragen, um nach diesem Zeitraum beide Gruppen miteinander vergleichen zu können. Es war den mitwirkenden Therapeuten jedoch nicht zwingend vorgegeben, die ‚Wartepatienten‘ auch tatsächlich ein Jahr warten zu lassen. Vielmehr sollte die Wartezeit sofort beendet sein, sobald der Therapeut den Wartepatienten nachrückend in Behandlung nehmen könnte.

Untersuchungs- und Kontrollgruppe sollten zu Beginn der Therapie bzw. zu Beginn der Wartezeit (definiert als Messzeitpunkt t1) mit dem MMPI(K) (Kurzform) und dem GBB befragt werden. Diese Befragung sollte nicht durch oder in Anwesenheit des Therapeuten, sondern sollte anonym erfolgen.

Hierzu bekamen die Therapeuten das mit einer Chiffre gekennzeichnete Testmaterial einschließlich eines Anschreibens an die Patienten, einer Instruktion und eines frankierten Rückumschlags für die Patienten zugeschickt. Die Therapeuten hatten lediglich die Aufgabe, das Testmaterial mit der Bitte um Bearbeitung an die Patienten auszuhändigen und um Weiterleitung der ausgefüllten Fragebögen an die Forschungsgruppe mit dem beiliegenden Freiumschlag zu bitten. (Die entsprechenden Anschreiben an die Therapeuten und die Patienten befinden sich im Anhang der Studie.)

Mit der ersten Datenerhebung zum Messzeitpunkt t1 wurden jeweils mit Hilfe des weiterleitenden Therapeuten die folgenden Daten erfasst: Name des Therapeuten, Patienten-Chiffre, Patientengeschlecht, Anzahl der bereits stattgefundenen Therapiesitzungen, Diagnose, Datum der Übergabe der Testunterlagen an den Patienten sowie Zugehörigkeit zur UG oder KG. Der Patient sollte dann noch das Datum der Beantwortung und seinen Schulabschluss nachtragen.

Nach Ablauf eines halben Jahres Datum der Bearbeitung der Testunterlagen bzw. nach Eingang der Testunterlagen bei der Forschungsgruppe wurden die Patienten über den gleichen Verfahrensweg gebeten, die gleichen Testunterlagen, zuzüglich des VEV, erneut zu bearbeiten (Testzeitpunkt t2) und wieder an die Forschungsgruppe zurückzuleiten. Hierbei war uns vor allem die Rückmeldung über die Anzahl der zwischenzeitlich stattgefundenen Therapiesitzungen wichtig. Dieser Vorgang wiederholte sich ein halbes Jahr später erneut (Testzeitpunkt t3).

Nach t3 wurden die Befragungsabstände auf ein volles Jahr ausgedehnt, d.h., t4 bezeichnet den Messzeitpunkt ein Jahr nach t3 und zwei Jahre nach t1. Der Messzeitpunkt t5 bezeichnet den Zeitraum nach einem weiteren Jahr, also zwei Jahre nach t3 und fünf Jahre nach t1. Die Kontrollgruppe bekam zwar zu den Zeitpunkten t4 und t5 ebenso Testunterlagen zugeschickt wie die Untersuchungsgruppe, jedoch wurden diese Daten dann im Rahmen der Kontrollgruppenzugehörigkeit aus Kostengründen später nicht mehr ausgewertet.

Was die Untersuchungsgruppendaten anbelangt, so wurden die Daten von Messzeitpunkt t4 ebenso aus Kostengründen nicht mehr ausgewertet, wohl aber die Daten von Messzeitpunkt t5.

Ab Sommer 2000 wurde die weitere Befragung der Patienten in den teilweise noch laufenden Therapien aus pragmatischen Begrenzungsgründen ausgesetzt und bei den Patienten mit zwischenzeitlich abgeschlossener Therapie wurde stattdessen eine katamnestische Nachbefragung mit Hilfe des Intrex-Kurzform-Fragebogens nach Tress, des Helping Alliance Questionnaire (HAQ) und einem selbst entwickelten ‚Fragebogen zur Freude in der Psychotherapie‘ durchgeführt. Einige Patienten hatten zu diesem Zeitpunkt die Therapie bereits seit Jahren, andere erst einigen Monaten beendet, - im Schnitt lag das Therapie-Ende ca. ein Jahr zurück. Die ‚Fragen zur Freude in der Therapie‘ lagen etwas außerhalb der Fragestellung dieser Studie. Der Hintergrund dieser Fragen liegt in einem über diese Studie hinausgehenden Forschungsinteresse der Autoren, welche Rolle die von der ‚ernsten Psychotherapie‘ tabuisierte ‚Freude‘ heutzutage spielt bzw. spielen darf (vgl. Kapitel 11.2.4 und Anhang).

Eine schematische Übersicht über das Studiendesign und den Ablaufplan der Untersuchung findet sich im gesonderten Anlageband zu dieser Studie.

## 7. Finanzierung

Überblick über die von der DGIP und den Alfred-Adler-Instituten zur Verfügung gestellten Forschungsmittel:

1. DGIP:	18.03.1993:	3.000,-- DM
2. DGIP:	27.05.1993:	5.000,-- DM
3. DGIP:	09.07.1993:	2.000,-- DM
4. AAI Düsseldorf:	16.08.1993:	5.000,-- DM
5. AAI Düsseldorf:	11.09.1993:	5.000,-- DM
6. DGIP:	22.07.1994:	10.000,-- DM
7. AAI Düsseldorf:	05.07.1995:	3.000,-- DM

8. AAI Düsseldorf:	11.10.1995:	2.000,-- DM
9. AAI München:	15.11.1995:	10.000,-- DM
10. AAI Delmenhorst:	23.12.1995:	5.000,-- DM
11. AAI Aachen:	01.03.1996:	3.000,-- DM
12. AAI Delmenhorst:	09.04.1998:	2.000,-- DM
Aus privaten Mitteln insgesamt	ca.	20.000,-- DM
Insgesamt also	ca.:	75.000,-- DM

## 8. Studiendurchführung

### 8.1 Organisation des Untersuchungsablaufs

Die Organisation des Untersuchungsablaufs erfolgte von Köln aus. Von hier aus wurden im April 1993 alle 230 Psychoanalytiker/-innen DGIP angeschrieben (vgl. Anlagenband) und um Mitwirkung gebeten. Auf Grund der schwachen Resonanz von nur 39 zur Mitarbeit bereiten Therapeuten/-innen wurde im Juni 1993 ein weiteres Motivationsschreiben verschickt (vgl. Anlagenband), woraufhin sich noch einmal 22 Therapeuten/-innen zur Mitarbeit bereit erklärten, wodurch sich die Gesamtzahl auf 61 erhöhte. Bezugnehmend auf die Gesamtzahl aller angeschriebenen Therapeuten entspricht dies einer Quote von 26,52 %. Später zeigte sich, dass fünf dieser zusagenden Therapeuten/-innen sich lediglich innerhalb der Kontrollgruppe mit einem Patienten, nicht jedoch in der Untersuchungsgruppe beteiligten. Letztendlich wirkten somit 55 Therapeuten mit unterschiedlich vielen Patienten in der Untersuchungsgruppe mit.

Parallel zu den „Werbeschreiben“ wurden die zur Mitarbeit bereiten Therapeuten/-innen mit einem erweiterten Schreiben (vgl. Anlagenband) über das Procedere, das Time-Management (Kontrolle der Zeitabläufe zur rechtzeitig wiederkehrenden Versendung der Fragebögen) durch das Forscherteam, das Anschreiben an die mitwirkenden Patienten, Untersuchungsdesign, Ablaufplan und den zu verwendenden Diagnoseschlüssel informiert. Darüber hinaus hefteten wir eine Stellungnahme zu den zwischenzeitlich von den Therapeuten geäußerten Fragen und Bedenken hinsichtlich der Bildung der Kontrollgruppe, der Anonymisierung, des Datenschutzes sowie zum einzelfallbezogenen Untersuchungsbeginn bei.

Im Zeitraum Juni 1993 – April 1994 kamen insgesamt 90 Patienten für die Untersuchungsgruppe und 10 Patienten für die Kontrollgruppe zur Anmeldung. Gleich mit den ersten Patientenmeldungen wurden auch bereits die vorbereiteten Fragebögen an die Therapeuten mit der Bitte um Chiffrierung und Weiterleitung an die Patienten verschickt (Messzeitpunkt t1). Diese Unterlagen an die Patienten waren mit frankierten Freiumschlägen zur Rücksendung an das Forscherteam versehen.

Mit Schreiben vom April 1994 wurden die mitwirkenden Therapeuten/-innen über den uns noch nicht genügenden Probandenbestand in Kenntnis gesetzt und um Einbringung weiterer Patienten gebeten. Im September 1994 erfolgte ein nochmaliges Anschreiben an die sich zur Mitarbeit bereiten Therapeuten mit der Bitte,

weitere Patienten in die UG und KG einzubringen. Insgesamt konnte die UG so auf 150 angemeldete Patienten, die KG auf 51 angemeldete Patienten gesteigert werden.

Im Jahre 2000 kam es schließlich zu einer Nachfabrikation zur Ermittlung von Schulabschluss und zum Beruf der Patienten bei Therapiebeginn.

## 8.2 Reaktion/Beteiligung der Therapeuten

Die Reaktion der angeschriebenen Therapeuten war sehr unterschiedlich. Einige Therapeuten/-innen waren sofort zur Mitarbeit bereit, andere reagierten überhaupt nicht, wieder andere verweigerten sich offen mit den verschiedensten Argumenten. Einige sahen ein quantitatives Studiendesign als nicht tauglich zur Untersuchung von Psychotherapien an; andere betonten ihre Unterstützung, brachten jedoch keine Patienten ein; wieder andere sahen es als ethisch nicht vertretbar an, eine Warteliste zu führen und diese Patienten zur Mitwirkung zu bewegen.

Von den ursprünglich 61 sich zur Mitarbeit bereit erklärenden Therapeuten zog eine Therapeutin ihre Zusage mit dem Argument zurück, die Patientin sei nicht zur Mitarbeit bereit gewesen, eine Therapeutin zog ihre sechs zuvor zum Messzeitpunkt t1 bereits eingebrachten Patienten wegen grundsätzlicher Auseinandersetzungen mit der DGIP, die mit dieser Untersuchung nichts zu tun hatten, zurück, vier Therapeuten brachten trotz vorher gegebener Zusage keine Patienten ein. Schließlich blieben 55 mitwirkende Therapeuten mit 144 in die UG und 51 in die KG eingebrachte Patienten übrig.

Diese verteilen sich wie folgt: Für die Untersuchungsgruppe fanden sich 19 Therapeuten, die mit je einem Patienten vertreten sind, 14 Therapeuten sind mit zwei Patienten vertreten, 13 Therapeuten mit drei Patienten, drei Therapeuten mit vier Patienten, drei Therapeuten mit sechs Patienten, zwei Therapeuten mit acht Patienten und ein Therapeutin mit zwölf Patienten, insgesamt also 55 Therapeuten mit 144 Patienten.

Für die Kontrollgruppe ergibt sich folgendes Bild: Von den 55 bereits in der UG mitwirkenden Therapeuten brachten 15 Therapeuten noch jeweils einen Patienten in die UG ein, acht Therapeuten waren mit jeweils zwei Patienten vertreten, fünf Therapeuten mit jeweils drei Patienten, ein Therapeut mit fünf Patienten, insgesamt also 28 Therapeuten mit 51 Patienten.

## 8.3 Stichprobenauswahl

Nach dem bisher Gesagten ist klar, dass sich die Auswahl der untersuchten Patienten im naturalistischen Sinne rein zufällig und nicht nach einem streng randomisierten Verfahren vollzog. Hiergegen kann mit Recht eingewendet werden, dass somit ein systematischer Fehler dergestalt mit eingeflossen sein könnte, als dass sich die Therapien der mitwirkungsbereiten Therapeuten u.U. bereits durch dieses Merkmal der „Mitwirkungsbereitschaft“ deutlich von den Therapien der „nicht mitwirkungsbereiten Therapeuten“ unterscheiden. Dies ist sicherlich nicht falsch, andererseits sahen wir keine Möglichkeit, zu einer statistisch „sauberen zufallsbedingten“ Auswahl zu kommen, so dass wir uns mit dieser naturalistischen Auswahl zufrieden gaben. Unsere Ergebnisse stehen somit allerdings unter dem methodologischen Vorbehalt einer möglichen Nachprüfung mittels einer Kontrolluntersuchung an Patienten aus einer nicht mitwirkungsbereiten Therapeutengruppe.

#### 8.4 Rücklaufquote

Die Rücklaufquote ist nicht identisch mit der Anzahl der von den Therapeuten zur Untersuchung angemeldeten Patienten. In einigen Fällen haben die Patienten die Therapie entweder nicht begonnen oder nach einigen Sitzungen abgebrochen, so dass diese Patienten dann wohl auch keine Motivation mehr zur Mitarbeit hatten und die zugesandten Unterlagen nicht zurücksandten. In anderen Fällen kam es sehr wohl zur Therapie, und dennoch schickten die Patienten die Unterlagen nicht zurück, was häufig erst zum Messzeitpunkt t2 oder t3 deutlich wurde. Wir hatten mit dieser Möglichkeit gerechnet, dass nämlich Patienten aus Gedankenlosigkeit oder Widerstand die Unterlagen nicht zurücksenden, wobei wir aus Gründen der Beibehaltung der Anonymität und um die Therapieprozesse nicht zu stören, beschlossen hatten, ausbleibenden Rücksendungen nicht nachzuspüren, - gleichwohl jedoch zum vorgesehenen Folgezeitpunkt trotzdem den Folgefragebogen zu verschicken und dann ggf. von dem vermittelnden Therapeuten zu erfahren, dass eine Weiterleitung aus welchen Gründen auch immer nicht möglich sei. In den Fällen, in denen die Therapie jedoch noch lief, kam es zur planmäßigen Weiterleitung und dann ggf. auch zur Rücksendung der ausgefüllten Fragebögen, obwohl die vorherigen Messzeitpunkte ohne Rückantwort geblieben waren.

So kommt es, dass zwischen der Anzahl der verschickten und der rückfließenden Fragebögen immer eine gewisse Differenz entstand. Zum Messzeitpunkt t1 wurden so z.B. 151 Fragebögen verschickt, zurück kamen jedoch lediglich 144 Fragebögen, wobei einige auch unausgefüllt oder nur teilausgefüllt zurück kamen, - wieder andere waren zwar weitgehend, aber dennoch immer wieder mit einigen unbeantwortet gebliebenen Items durchsetzt.

Nicht alle dieser verwertbaren Fragebögen kamen jedoch letztlich zur Auswertung, was damit zu tun hat, dass einige der Patienten zum Zeitpunkt t3 bereits ausgeschieden (Therapieabbruch oder Ähnliches) waren, so dass deren Vergleichsdaten zum Messzeitpunkt t3 nicht mehr erhoben werden konnten, was zur Folge hatte, dass in die spätere Auswertung des Vergleichszeitraums Messzeitpunkt t1 zu Messzeitpunkt t3 nur noch die Patienten hineinkamen, die sowohl zum Zeitpunkt t1 wie auch zum Zeitpunkt t3 komplette und verwertbare Daten geliefert hatten. So kamen zwar zum Messzeitpunkt t1 insgesamt 144 verwertbare Datensätze zurück, - da jedoch von diesen 144 Patienten zum Messzeitpunkt t3 nur noch 110 Patienten durch verwertbare Rückläufe vertreten waren, schieden 34 Patienten trotz verwertbarer Daten zum Zeitpunkt t1 aus der weiteren Untersuchung aus und das „N“ für die erste Vergleichsauswertung der Messzeitpunkte t1 zu t3 sank, - festgestellt auf der Basis vollständiger Daten zur GBB-Skala „Erschöpfungsneigung“ zum Messzeitpunkt t3 -, auf 110. Eine komplette Übersicht über den Verlauf der Teilnehmerzahl in der UG gibt Kap. 10.1 (vgl. unten).

Die KG begann zum Messzeitpunkt t1 mit 51 wartenden Patienten. Ein halbes Jahr später, also zum Messzeitpunkt t2, erfolgte ein Rücklauf der versandten Fragebögen von noch 41, zum Messzeitpunkt t3 von noch 39 wartenden Patienten.

### 8.5 Datenqualität

Die Datenqualität war nicht immer gut. In einzelnen Fällen kamen komplette Datensätze von den Patienten unausgefüllt zurück. In den meisten Fällen waren die Datensätze zwar ausgefüllt, jedoch kam es immer wieder vor, dass einzelne Items nicht beantwortet waren. Über die tatsächlich verwertbaren Rücklaufquoten geben die Auswertungstabellen zum UG/KG-Gruppenvergleich bzw. die Auswertungen zur Varianzanalyse Auskunft, wobei die Höhe der „N“ hier schwankten, je nach dem, wie viele einzelne Items einer Skala trotz ausgefülltem Gesamtbogen im individuellen Einzelfall dennoch unbeantwortet geblieben waren. Der „verwertbare“ Rücklauf zum Messzeitpunkt t3 liegt bei ca. 91-109 („N“), zum Zeitpunkt t5 bei ca. 54-65 („N“). Auf Grund der Zielsetzung, die Anonymität der Patienten auf keinen Fall aufheben zu wollen, aber auch weil ansonsten ein nicht zu bewältigender Nachfassaufwand hätte betrieben werden müssen, wurde das eingehende Datenmaterial nicht durch Rückfragen bei den Patienten vervollständigt.

### 8.6 Doppelte Dateneingabe und stichprobenartige Überprüfung

Alle Daten wurden von einer unbeteiligten und honorierten Hilfskraft gemäß einem vorher festgelegten Schlüssel im Programm „Word“ verschlüsselt eingegeben. Hierbei wurden zunächst folgende Stammdaten der Reihe nach in codierter Form eingegeben: Satz-Nummer, Patientennummer, Therapeuten-Chiffre, Patienten-Chiffre, Patientengeschlecht, Anzahl der Therapiesitzungen zum Messzeitpunkt, Diagnose-Schlüssel, Datum der Übergabe, UG/KG, Datum der Beantwortung, Messzeitpunkt, Postleitzahl des Therapeuten, Geschlecht des Therapeuten, Beruf des Therapeuten, Jahr der Prüfung des Therapeuten, Ort des Ausbildungsinstituts des Therapeuten, Datum der Therapiebeendigung, Grund der Beendigung, Chiffre des Datenerfassers. Im Anschluss daran erfolgte die codierte Eingabe aller Itemantworten oder Fehlantworten gemäß GBB, MMPI(K) und VEV, für alle Patienten und für alle Messzeitpunkte.

Zur Überprüfung der Eingabequalität wurden stichprobenartig Datensätze von einer anderen unbeteiligten Hilfskraft doppelt eingegeben und mit der Eingabe der ersten Hilfskraft verglichen. Hierbei stellte sich heraus, dass die Fehlerquote der eingegebenen Daten weit unter 1% lag und gegen Null tendierte.

## 9. Grundsätze bei der Hypothesenformulierung

Im Sinne des Popper'schen Wissenschaftsverständnisses des so bezeichneten „Kritischen Rationalismus“ gibt es keine naiv-positivistische Beobachtung, sondern immer nur eine theorie- oder hypothesengeleitete Beobachtung. In diesem Punkt sind sich die ansonsten wissenschaftstheoretisch doch so verschiedenen Autoren Popper und Adorno einig. Dahrendorf führt hierzu aus:

„Beide bezeichneten sich ... als Negativisten, insofern sie die Aufgabe der Empirie in der kritischen Korrektur sahen. Darüber hinaus betonten beide immer wieder den Primat der Theorie in der Wissenschaft. Für Popper ergibt sich dieses Primat aus der eindeutigen Verknüpfung von Theorie und Empirie in der ‚hypothetisch-deduktiven Methode‘ der Wissenschaft ...: Es gibt keine Beobachtung ohne Hypo-

these ... Die Induktion ist die falsche These, dass man von der Beobachtung ausgehen kann. Induktion gibt es eben nicht.“ (Dahrendorf, 1972, S. 148)

Wenn also jedes wissenschaftliche Vorgehen nie von einer naiven Beobachtung, sondern immer nur von einer Hypothese ausgehen kann, dann heißt dies für die Untersuchung, dass es zunächst gilt, Hypothesen zu formulieren.

Stellt sich die Frage, wie diese Hypothesen wissenschaftlich exakt formuliert werden müssen. Hierzu führt Habermas wiederum Popper an:

„Bekanntlich führt er (Popper) den Nachweis, dass Gesetzhypothesen überhaupt nicht verifiziert werden können. Diese Hypothesen haben die Gestalt unbeschränkter Allsätze mit einer unbegrenzten Zahl prinzipiell möglicher Anwendungsfälle, während ja die Reihe der Beobachtungen, mit deren Hilfe wir jeweils an einem Fall die Hypothese überprüfen, ebenso prinzipiell endlich ist. Ein induktiver Beweis ist daher unmöglich. Gesetzhypothesen lassen sich allenfalls indirekt dadurch bestätigen, dass sie möglichst vielen Versuchen der Falsifikation standhalten. Scheitern kann eine Theorie an singulären Existenzbehauptungen, die der in eine negative Voraussage umformulierten Gesetzhypothese widersprechen.“ (Habermas, 1972, S. 176)

Dies bedeutet, dass nicht die eigentlich interessierende Annahme über einen in der Grundgesamtheit existierenden Zusammenhang als Arbeitshypothese gewählt wird, sondern die Annahme, die es zu widerlegen gilt. Dies heißt streng genommen für unsere Untersuchung, dass wir unsere Hypothesen eigentlich immer in der Art „negativer Voraussagen“ formulieren müssten. Wenn wir anders verfahren, nämlich einen vermuteten und deshalb „positiv formulierten hypothetischen Zusammenhang“ durch Beibringung eines durch unsere Untersuchung zu Tage geförderten Belegs induktiv zu beweisen suchen, dann bewegen wir uns damit streng genommen nicht mehr auf dem Boden der Wissenschaftstheorie des Kritischen Rationalismus nach Popper. Gleichwohl nehmen wir diese kleine wissenschaftstheoretische Unkorrektheit in der Formulierung wissentlich in Kauf und begründen dies damit, dass die „negative Umformulierungen“ die meisten Leser verwirren und deshalb in der Literatur aus Gründen der „Leserfreundlichkeit“ auch nur selten vorzufinden sind. Lediglich in den Kapiteln 10.4.6 – 10.4.11 haben wir uns aus exemplarischen Gründen bemüht, wissenschaftstheoretisch korrekt negativ umformulierte Hypothesen derart zu formulieren, dass wir nämlich vermuten, dass zwischen den unabhängigen und abhängigen Variablen kein Zusammenhang besteht. Wir formulieren die negativ umformulierten Hypothesen als Ho-Hypothesen, wobei mit Nullhypothese üblicherweise die zu testende Annahme über die Grundgesamtheit bezeichnet wird. Und auch nachdem negativ umformulierte Ho-Hypothesen falsifiziert sind, ist es im Sinne der wissenschaftstheoretischen Position des „Kritischen Rationalismus“ wissenschaftstheoretisch unzulässig, daraus im Umkehrschluss folgernd einen „verifizierten“ Zusammenhang zu behaupten. Dies deshalb, weil es ja nur eines einzigen dem widersprechenden singulären Ereignisses bedarf, um einen verfrüht als verifiziert angenommenen Satz zu falsifizieren. Statthaft hingegen ist es, eine statistische Wahrscheinlichkeit (Signifikanzniveau) für die Existenz eines Zusammenhangs zu



benennen, - wodurch der Falsifikationsmöglichkeit eine auch über ein einmalig erzielttes Forschungsergebnis hinaus weiterhin existierende statistische Existenzzuordnung zugestanden wird.

Statistische Wahrscheinlichkeitsrechnung und der „Kritische Rationalismus“ Poppers gehen hier eine geradezu notwendige wissenschaftstheoretische Allianz ein.

## 10. Ergebnisse der Effektivitätsstudie (Verlaufsstudie)

In den folgenden Kapiteln wollen wir die Ergebnisse unserer Verlaufsstudie darstellen. Beginnen wollen wir mit der Darstellung der Entwicklung unserer Teilnehmerzahlen (Kap. 10.1). Es folgt eine erste allgemeine Übersicht über die zu den verschiedenen Messzeitpunkten auf den klinischen Skalen unserer Messverfahren in der Untersuchungs- und Kontrollgruppe gemessenen Mittelwerte (Kap. 10.2). Der direkte deskriptive Vergleich dieser Werte gibt einen ersten Eindruck über die Ergebnisse der Verlaufsstudie. Daran schließt sich die statistische Verarbeitung und Prüfung dieser Werte hinsichtlich verschiedener Ergebnismaße für Therapieeffektivität an (Kap. 10.3). Danach kommen wir zur Darstellung der Ergebnisse von Varianzanalysen, die wir hinsichtlich von 11 die Untersuchungsgruppe dichotom beschreibenden unabhängigen Variablen durchgeführt haben (Kap. 10.4). Das Kapitel 10.5 schließlich fasst die Ergebnisse zusammen.

### 10.1 Die Anzahl der Teilnehmer zu den fünf Testzeitpunkten

Die Untersuchung startete zum Messzeitpunkt t1 in der Untersuchungsgruppe mit 144 Patienten, in der Kontrollgruppe mit 51 Patienten (gemessen auf der Basis vollständiger Daten in GBB-1-Erschöpfungsneigung). Es war zu erwarten, dass sich diese Zahl bis zum Messzeitpunkt t5, also drei Jahre nach Behandlungsbeginn, verkleinern würde.

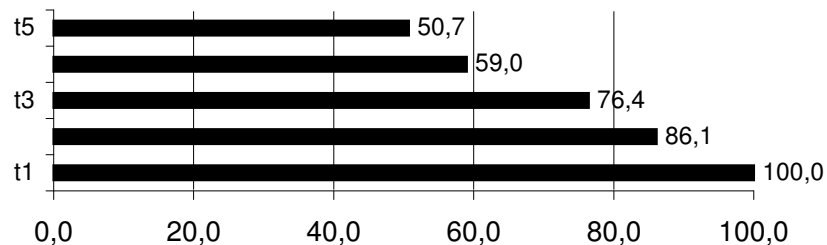
Tabelle 2 zeigt den Rückgang der Teilnehmerzahl im Zeitraum von t1 (Zeitpunkt der probatorischen Sitzungen) bis zu t5 (nach drei Jahren Behandlung). Die Teilnehmerzahl hat sich nahezu halbiert. Der prozentual stärkste Rückgang erfolgte im ersten Behandlungsjahr (Messzeitpunkt t1 bis Messzeitpunkt t3) und erreichte dort 23,6 %-Punkte.

**Tabelle 2: Teilnehmerzahl in der Untersuchungsgruppe und der Kontrollgruppe zu den jeweiligen Messzeitpunkten auf der Basis vollständiger Daten zur GBB-Skala-1 „Erschöpfungsneigung“**

	t1	t2	t3	t4	t5
N/UG	144	124	110	85	73
N/KG	51	41	39	-	-

Eine graphische Darstellung dieser Entwicklung in der Untersuchungsgruppe zeigt Abbildung 3:

**Abb. 3: Teilnehmerzahl in der Untersuchungsgruppe zu den fünf Messzeitpunkten (in Prozent)**



In der Kontrollgruppe ergab sich ein Rückgang der Anzahl der Patienten von 51 zum Messzeitpunkt t1 über 41 zum Messzeitpunkt t2 bis hin zu 39 zum Messzeitpunkt t3, dem Zeitpunkt der Auflösung der Kontrollgruppe.

Der von uns in beiden Gruppen festgestellte Rückgang der Teilnehmerzahl deckt sich weitgehend mit den Erfahrungen von Tschuschke und Anbeh (2000), die auf Grund nicht vorhandener Compliance bei solchen Untersuchungen von einer Dropout-Rate in Höhe von bis zu 50 % ausgehen. Bei ihnen nahmen zu Beginn der Gruppenbehandlung zunächst 545 Patienten teil, - vollständige Daten zu allen drei Messzeitpunkten konnten jedoch nur von 194 Patienten/innen erhoben werden.

Kächele (1990) berichtet zu der Frage der Dauer von Psychotherapien über Ergebnisse aus einer universitären psychotherapeutischen Ambulanz: Die analytische Psychotherapie, zweistündig im Sitzen durchgeführt, sei überraschenderweise bei über 60 % der Patienten bereits nach den ersten beantragten 80 Sitzungen beendet gewesen. Hier seien jedoch zu einem erheblichen Teil Veränderungen der Therapieziele und damit auch Veränderungen der Planung der Behandlung maßgeblich gewesen; einseitig vom Patienten initiierte Therapieabbrüche machen in dieser Gruppe nicht mehr als 10 % aus.

Die Analyse von Behandlungsabbrüchen in der stationären psychosomatischen Rehabilitation (Schulz et al., 1999) ergab in den untersuchten Stichproben Quoten vorzeitiger Entlassungen von 8,3 % bzw. 14,7%. In der Literatur sind Berichte über Bandbreiten von 10,6 % bis 30,1 % Abbruchquote anzutreffen.

## 10.2 Überblick über Mittelwerte und Standardabweichungen in den GBB-Skalen, den MMPI(K)-Skalen und im VEV der Untersuchungsgruppe (UG) und der Kontrollgruppe (KG) zu den verschiedenen Messzeitpunkten

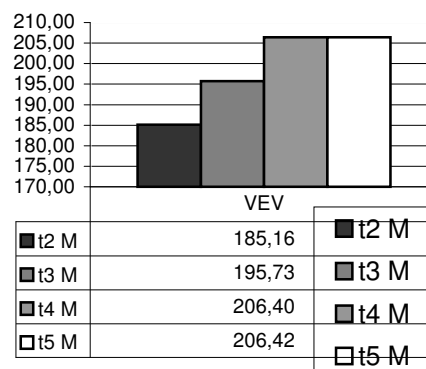
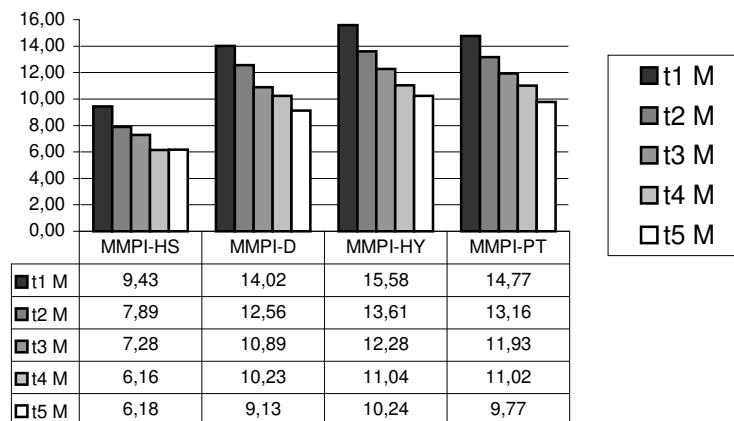
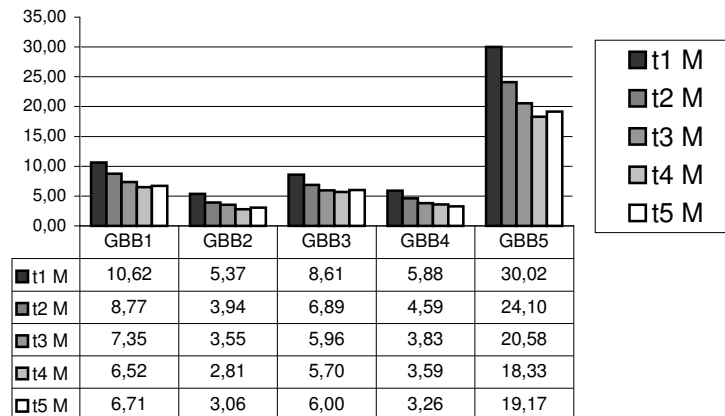
Tabelle 3 gibt eine Übersicht über die Anzahl der Patienten, die Mittelwerte und die Standardabweichungen in den fünf Skalen des GBB, den 13 Skalen des MMPI(K) und im VEV zu jedem der fünf Testzeitpunkte in der Untersuchungsgruppe. Von t1 bis t5 ist ein Rückgang der Teilnehmerzahl um ca. die Hälfte feststellbar. Die gemessenen Skalenmittelwerte zeigen bei den GBB-Skalen eine abnehmende Tendenz bis zum Zeitpunkt t4, sinken dann aber bis zum Messzeitpunkt t5 nicht weiter ab. Bei den MMPI(K)-Skalen bleiben die beiden Kontrollskalen L und F relativ konstant, während die Mittelwerte der Skala K von Messzeitpunkt t1 bis Messzeitpunkt t5 ansteigen. Bei den für diese Studie besonders interessanten Skalen der neurotischen Trias (Hypochondriasis (HS), Depression (D), Hysterie (HY)) sowie Psychasthenie (PT) ist von t1 bis t5 (bei HS bis t4) eine mehr oder weniger kontinuierliche Abnahme der Mittelwerte zu verzeichnen, ebenso in der Skala Psychopathie (PP) und Schizoidie (SC) bis zum Messzeitpunkt t4. Die Mittelwerte der M/F-Interessen-Skala bleiben über die gesamte Zeit unverändert.

**Tabelle 3: Anzahl vollständiger Antworten (N), Mittelwerte (M) und Standardabweichungen (s) in den GBB-Skalen, den MMPI(K)-Skalen und im VEV der Untersuchungsgruppe zum Zeitpunkt der probatorischen Sitzungen (t1) nach einem halben Jahr (t2), nach einem Jahr (t3), nach zwei Jahren (t4) und nach drei Jahren Behandlung (t5)**

	t1			t2			t3			t4			t5		
	N	M	s	N	M	s	N	M	s	N	M	s	N	M	s
GBB1-Erschöpfungsneigung	144	10,62	5,86	124	8,77	5,74	110	7,35	5,26	85	6,52	5,01	73	6,71	5,06
GBB2-Magenbeschwerden	143	5,37	4,44	124	3,94	3,56	110	3,55	3,40	83	2,81	3,16	71	3,06	3,29
GBB3-Gliederschmerzen	141	8,61	5,40	122	6,89	4,58	110	5,96	4,14	84	5,70	4,62	73	6,00	4,59
GBB4-Herzbeschwerden	141	5,88	3,99	124	4,59	3,65	109	3,83	3,80	83	3,59	3,62	73	3,26	3,42
GBB5-Beschwerdedruck	134	30,02	14,87	122	24,10	13,26	106	20,58	12,40	83	18,33	12,58	71	19,17	13,14
MMPI-Lie	142	3,45	1,95	116	3,57	1,73	100	3,80	1,81	82	3,32	1,80	68	3,65	1,67
MMPI-Validity	139	5,00	3,87	117	4,39	2,92	94	4,31	3,29	75	3,81	2,67	65	3,48	2,62
MMPI-Correction	140	8,50	3,36	121	9,08	3,65	109	9,70	3,74	83	9,73	3,76	69	10,41	3,80
MMPI-Hypochondriasis	141	9,43	4,70	123	7,89	4,39	109	7,28	4,50	85	6,16	4,47	71	6,18	4,56
MMPI-Depression	144	14,02	4,38	122	12,56	4,70	108	10,89	5,12	81	10,23	4,91	72	9,13	5,50
MMPI-Hysteria	137	15,58	5,39	121	13,61	4,93	104	12,28	5,90	80	11,04	5,41	70	10,24	5,79
MMPI-Psychopathie	136	12,75	4,41	118	12,06	4,52	103	10,56	4,55	76	9,78	4,42	70	9,27	4,45
MMPI-maskulin-feminin	133	15,00	3,39	121	14,82	3,54	109	14,73	3,10	80	14,50	3,75	70	14,47	3,69
MMPI-Paranoia	134	8,17	3,94	120	9,20	3,51	104	6,78	3,65	82	6,02	3,17	71	5,76	3,20
MMPI-Psychasthenia	140	14,77	5,11	121	13,16	5,26	109	11,93	5,58	82	11,02	5,79	71	9,77	5,50
MMPI-Schizoidie	133	9,42	4,86	117	8,30	4,60	105	7,28	4,68	79	5,94	4,13	61	6,07	3,82
MMPI-Hypomanie	134	10,40	3,16	121	10,19	3,46	104	9,90	3,67	81	9,73	4,10	68	9,47	3,69
MMPI-Soziale Introversion	141	13,83	4,59	121	14,18	4,67	109	12,93	4,71	80	13,00	5,15	68	12,07	5,49
VEV				124	185,16	28,31	108	195,73	37,06	86	206,40	34,41	72	206,42	43,76

Eine graphische Darstellung dieser Ergebnisse zeigt die Abbildung 4:

**Abb. 4: Mittelwerte der Messzeitpunkte t1 bis t5, gemessen in den fünf GBB-Skalen, den 4 MMPI(K)-Skalen (neurotische Trias und PT) und im VEV der Untersuchungsgruppe**



In der nachfolgenden Tabelle 4 sind alle Häufigkeiten, Mittelwerte und Standardabweichungen in den GBB-Skalen, MMPI(K)-Skalen und im VEV für die Kontrollgruppe aufgeführt. Bei einem ersten Überblick über die Daten zeigt sich, dass sich die Mittelwerte und Standardabweichungen für die einzelnen Skalen zu t1, t2 und t3 nur unwesentlich unterscheiden.

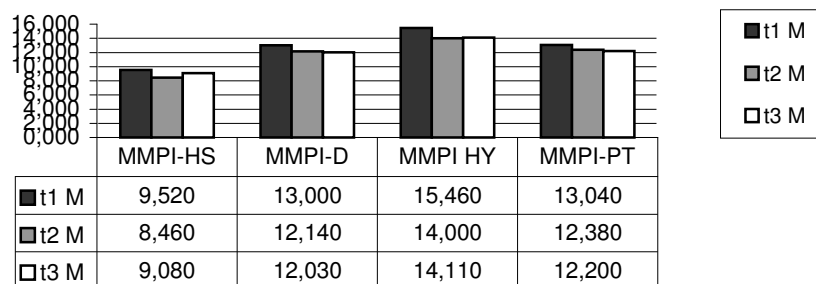
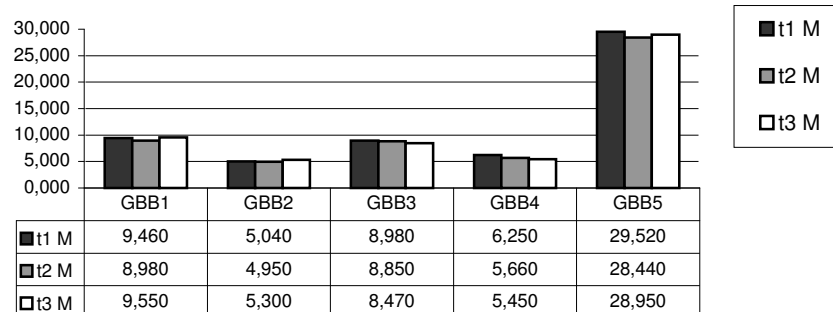
**Tabelle 4: Mittelwerte (M) und Standardabweichungen (s) in den GBB-Skalen, den MMPI(K)-Skalen und im VEV der Kontrollgruppe zum Zeitpunkt der probatorischen Sitzungen (t1), nach einem halben Jahr Wartezeit (t2) und nach einem Jahr Wartezeit (t3)**

	t1			t2			t3		
	N	M	s	N	M	s	N	M	s
GBB1-Erschöpfungsneigung	50	9,460	5,760	41	8,980	5,510	38	9,550	6,470
GBB2-Magenbeschwerden	51	5,040	3,700	41	4,950	4,350	37	5,300	4,130
GBB3-Gliederschmerzen	51	8,980	5,700	41	8,850	5,530	38	8,470	6,030
GBB4-Herzbeschwerden	51	6,250	4,390	41	5,660	4,840	38	5,450	5,510
GBB5-Beschwerdedruck	50	29,520	16,600	41	28,440	17,350	37	28,950	18,650
MMPI-Lie	50	3,740	1,800	40	3,680	1,610	37	3,510	1,730
MMPI-Validity	45	4,380	3,610	32	4,660	2,910	32	4,660	3,580
MMPI-Correction	45	8,290	3,540	35	8,800	3,900	36	8,610	3,560
MMPI-Hypochondriasis	50	9,520	4,530	39	8,460	5,210	38	9,080	5,380
MMPI-Depression	49	13,000	5,390	37	12,140	5,410	37	12,030	6,130
MMPI-Hysteria	48	15,460	5,730	37	14,000	5,950	37	14,110	6,790
MMPI-Psychopathie	42	10,780	5,180	37	10,730	5,360	36	10,750	5,880
MMPI-maskulin-feminin	49	14,260	3,280	39	13,670	3,290	37	13,680	3,980
MMPI-Paranoia	48	7,290	4,000	36	6,970	4,620	36	6,970	4,130
MMPI-Psychasthenia	49	13,040	6,070	40	12,380	6,480	35	12,200	6,630
MMPI-Schizoidie	47	8,380	5,410	36	8,420	4,880	33	9,030	5,210
MMPI-Hypomanie	46	10,890	3,830	33	10,090	3,890	33	10,760	4,620
MMPI-Soziale Introversion	49	12,860	5,720	36	13,640	5,250	37	12,320	5,780
VEV	-	-	-	41	171,560	32,460	39	173,180	38,830

Sowohl die Tabelle 3 für die Untersuchungsgruppe wie auch Tabelle 4 für die Kontrollgruppe zeigen, dass das „N“ zu allen drei Messzeitpunkten zwischen den Skalen des MMPI(K) variiert, was darauf zurückzuführen ist, dass die Patienten zu einigen Items keine oder nicht auswertbare Antworten gegeben haben. Auffallend ist, dass es bei den Skalen F und K relativ viele nicht auswertbare Antworten gibt, bei den vier MMPI(K)-Skalen der neurotischen Trias (Hypochondriasis, Depression, Hysterie) und Psychasthenie dagegen relativ wenige Fehlantworten. Die Fragen des VEV wurden anweisungsgerecht komplett beantwortet.

Eine graphische Darstellung der Ergebnisse in Tabelle 4 zeigt Abbildung 5:

**Abb. 5: Mittelwerte der Messzeitpunkte t1 bis t3, gemessen in den fünf GBB-Skalen, den vier MMPI(K)-Skalen (neurotische Trias + PT) der Kontrollgruppe**



Zentrales Anliegen unsere Studie ist die Beantwortung der Frage, ob sich mit den dargestellten Mittelwerten eine signifikante Wirksamkeit der psychotherapeutischen Behandlung nachweisen lässt und in welchem Ausmaß und Phasen diese ggf. verläuft. Um diese Frage zu beantworten, werden die dargestellten Messwerte nachfolgend mittels statistischer Prüfverfahren auf ihre Aussagefähigkeit hin untersucht.

### 10.3 Sechs Ergebnismaße für die Therapieeffektivität

Wir prüfen den Therapieerfolg mittels fünf verschiedener Ergebnismaße, als da wären:

10.3.1: Die Signifikanzprüfung der direkten Einschätzung der Veränderung des Erlebens und Verhaltens durch die Patienten mittels des Veränderungsfragebogens des Erlebens und Verhaltens (VEV) von M. Zielke und C. Kopf-Mehnert.

10.3.2: Die Signifikanzprüfung der Mittelwertvergleiche der in den Skalen des GBB, des MMPI(K) und des VEV erzielten Messergebnisse der behandelten Untersuchungs- und der unbehandelten Kontrollgruppe zu den Messzeitpunkten t1, t2 und t3 (Querschnittvergleiche).

10.3.3: Die Signifikanzprüfung der Prä-Post-Vergleiche der Messergebnisse in den Skalen des GBB, des MMPI(K) und des VEV sowohl der Untersuchungs- wie auch der Kontrollgruppe zwischen den verschiedenen Messzeitpunkten der jeweiligen Gruppe (Längsschnittvergleiche).

10.3.4: Varianzanalyse zur integrierten Signifikanzprüfung des Entwicklungsvergleichs von UG und KG.

10.3.5: Die indirekte Erfolgsmessung durch Vergleich unserer Messergebnisse in den GBB-Skalen, den MMPI(K)-Skalen sowie im VEV mit den Normwerten externer Vergleichsgruppen, so wie sie im Zusammenhang mit der Eichung der von uns verwendeten Messinstrumente in der Literatur veröffentlicht sind.

10.3.6: Die Berechnung des Therapieeffektes (Effektstärke) zu verschiedenen Zeitpunkten mit Hilfe des „Globalmaßes der Therapieeffektivität“.

### 10.3.1 Ergebnisse im VEV zu den Testzeitpunkten t2, t3, t4 und t5

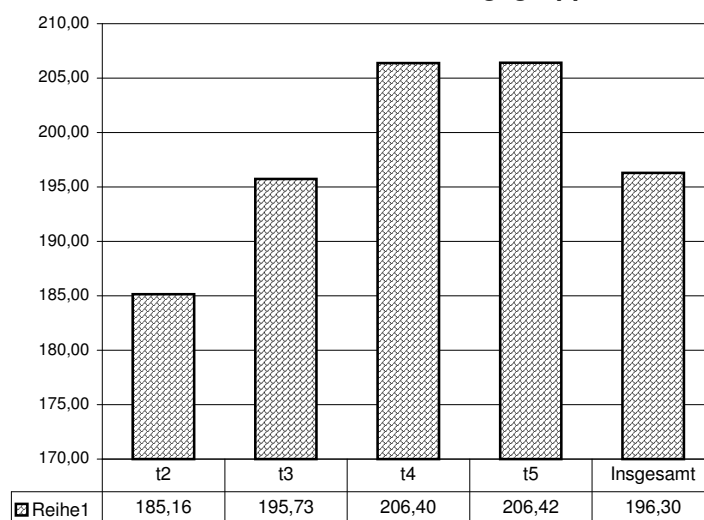
Tabelle 5 zeigt die Entwicklung der Mittelwerte im VEV von t2 bis t5. Der Mittelwert aller über vier Messzeitpunkte erhobenen VEV-Werte beträgt 196,30 und liegt damit fast exakt bei dem Mittelwert zum Messzeitpunkt t3.

**Tabelle 5: Der VEV-Gesamtwert von t2 bis t5 in der Untersuchungsgruppe**

	t2	t3	t4	t5	Insgesamt
<b>Mittelwert</b>	185,16	195,73	206,40	206,42	196,30
<b>N</b>	124	108	86	72	391
<b>Standardabweichung</b>	28,31	37,06	34,41	43,67	37,08

Eine graphische Darstellung der Tabelle 5 zeigt Abbildung 6:

**Abb. 6: VEV-Mittelwerte von Messzeitpunkt t2 bis Messzeitpunkt t5 in der Untersuchungsgruppe**



Die folgende Tabelle 6 liefert einen Überblick über die kritischen Differenzen der relevanten VEV-Maße:



**Tabelle 6: Kritische Differenzen zum Nachweis statistisch bedeutsamer Veränderungen im VEV (Zielke, 1978)**

				kritische Spanne
Entspannung	{	0,1 % *	Verbesserung	200-294
Gelassenheit		1,0 % *	Verbesserung	193-199
Optimismus		5,0 % *	Verbesserung	187-192
Keine Änderung				
Spannung	{	5,0 % *	Verschlechterung	144-149
Unsicherheit		1,0 % *	Verschlechterung	137-143
Pessimismus		0,1 % *	Verschlechterung	42-136
* Signifikanzniveau p = oder < als				

Nach den Kriterien der „kritischen Differenzen“ lässt sich zum Zeitpunkt t2 noch von keiner signifikanten Veränderung sprechen. Bis zum Zeitpunkt t3 steigt der Mittelwert um zehn Punkte, bis t4 um weitere elf Punkte, um dann bis t5 auf diesem Niveau zu stagnieren. Der Mittelwert zu t3 zeigt auf dem Signifikanzniveau von 1 %, der Mittelwert zu t4 auf dem Signifikanzniveau von 0,1 % eine signifikante Verbesserung im Vergleich zum Therapiebeginn.

Die Standardabweichung des Mittelwertes zu t5 ist vergleichsweise hoch, was auf eine heterogene Entwicklung hinsichtlich Verbesserung und Verschlechterung hinweist.

Zur Interpretation des Befundes zu t5: Möglicherweise zeigen sich in diesem Ergebnis die Schwankungen im Verlauf von Langzeitanalysen, von denen auch Sandell (2001) spricht. Die in unserer UG zu t5 befindlichen Patienten sind ja noch nicht am Ende der Behandlung angekommen, d.h., mit dem Veränderungsmaß zu t5 ist noch nichts über den Behandlungserfolg zum oder nach Abschluss der Behandlung ausgesagt. Vielleicht ist es aber auch so, dass es trotz Langzeittherapie nicht bei allen Patienten gelingt, Fortschritte über ein bestimmtes Maß hinaus zu erzielen; hier bliebe zu untersuchen, ob dies im Zusammenhang mit bestimmten Störungsbildern oder auch mit bestimmten Therapeutenvariablen, z.B. mit dessen Eignung für ein mehr hochfrequentes psychoanalytisches oder für ein mehr niederfrequentes tiefenpsychologisch fundiertes Vorgehen (s. Sandell, 2001), steht.

### **10.3.2 Einschätzung der Veränderungen in der Untersuchungsgruppe durch Vergleich mit der Kontrollgruppe unbehandelter Patienten im und nach dem ersten Behandlungsjahr (Querschnittuntersuchung)**

Um überhaupt die Behauptung aufstellen zu können, dass ein eventuell signifikanter Unterschied zwischen UG und KG auch tatsächlich nur auf die intervenierende Variable (psychotherapeutische Behandlung) zurückzuführen ist, muss als Voraussetzung die Strukturgleichheit beider Gruppen gegeben

sein, - was zunächst zu prüfen ist. Dazu werden nachfolgend die Untersuchungsgruppe und die Kontrollgruppe in ihrer Zusammensetzung hinsichtlich verschiedener Merkmale miteinander verglichen.

### 10.3.2.1 Prüfung der Strukturgleichheit von Untersuchungsgruppe (UG) und Kontrollgruppe (KG) als notwendige Voraussetzung des Gruppenvergleichs

Zur Überprüfung der Strukturgleichheit von Untersuchungs- und Kontrollgruppe ziehen wir folgende Merkmale heran: Alter, Geschlecht, Schulabschluss, Beruf, Diagnosestellung (zu t1), Diagnosegruppen (lt. ICD 10) sowie der Sitz des Ausbildungsinstituts der Therapeuten.

a) Prüfung der Strukturgleichheit der Gruppen hinsichtlich des Alters der Gruppenmitglieder:

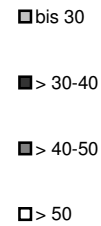
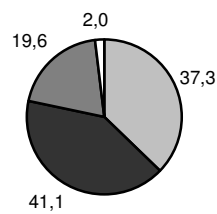
**Tabelle 7.1: Zusammensetzung von Untersuchungsgruppe und Kontrollgruppe nach Altersgruppen, erhoben zum Messzeitpunkt t1**

	UG/ t1		KG/ t1	
	N	Prozent	N	Prozent
<b>bis 30</b>	43	30,1	19	37,3
<b>&gt; 30-40</b>	73	51,0	21	41,1
<b>&gt; 40-50</b>	23	16,1	10	19,6
<b>&gt; 50</b>	4	2,8	1	2,0
<b>Gesamt</b>	143	100,0	51	100,0

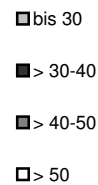
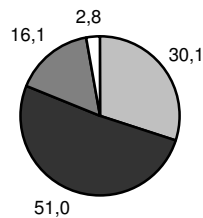
Durchschnittsalter		
	UG	KG
<b>M</b>	34,14	34,67

Tabelle 7.1 weist sowohl in der Untersuchungsgruppe als auch in der Kontrollgruppe ein Überwiegen der Altersgruppe der 30- bis 40-jährigen Patienten aus, wobei diese Altersgruppe in der Untersuchungsgruppe um 10 % stärker als in der Kontrollgruppe vertreten ist. Die Altersklasse der bis 30-jährigen Patienten stellt in beiden Gruppen die zweitstärkste Fraktion, ist allerdings in der KG ca. 7 % stärker vertreten als in der UG. Die drittstärkste Altersklasse ist die der 40- bis 50-Jährigen, die viertstärkste Klasse die der über 50-jährigen, wobei diese beiden Altersklassen in UG und KG in etwa gleich stark/schwach vertreten sind. Insgesamt gesehen möchten wir unsere UG und KG, - bezogen auf das Alter -, als relativ strukturgleich bezeichnen, zumal das Durchschnittsalter in beiden Gruppen bei einem nahezu identischen Wert von ca. 34 Jahren liegt. Eine graphische Darstellung der Zusammensetzung von UG und KG nach Altersgruppen zeigen die Abbildungen 7 und 8:

**Abb. 8: Zusammensetzung der Kontrollgruppe nach Altersgruppen (in Prozent), erhoben zum Messzeitpunkt t1**



**Abb. 7: Zusammensetzung der Untersuchungsgruppe nach Altersgruppen (in Prozent), erhoben zum Messzeitpunkt t1**



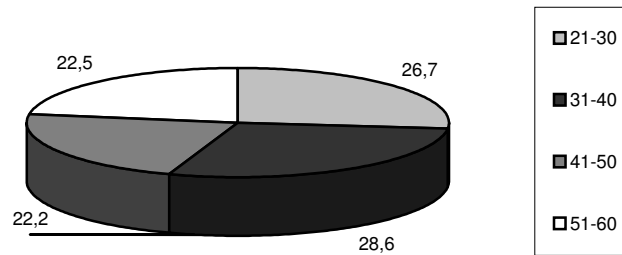
Wir vergleichen die Altersstruktur unserer UG und KG mit der Altersverteilung der bundesrepublikanischen Bevölkerung:

**Tabelle 8: Bevölkerung (BRD), 1994, nach Altersgruppen**  
(Quelle: Statistisches Bundesamt)

Altersgruppe	N	%
21-30	11.104.261	26,7
31-40	11.893.630	28,6
41-50	9.235.456	22,2
51-60	9.342.174	22,5
<b>Gesamt (21-60)</b>	<b>41.575.521</b>	<b>100,0</b>

Eine graphische Darstellung der Tabelle 8 zeigt Abbildung 9:

**Abb. 9: Bevölkerung in Prozent nach Altersgruppen**



Die Bundesrepublikanische Bevölkerung hat keinen wirklichen Schwerpunkt in einer der aufgeführten Altersklassen. Mit relativ geringer Abweichung nach oben und nach unten sind alle Altersgruppen mit Werten um die 25 % besetzt. Die Besetzung unserer UG mit einem Schwerpunkt von 51 % in der Altersklasse der 30- bis 40-Jährigen und unserer KG mit 41 % in der gleichen Altersklasse weicht deutlich von der Altersstruktur der BRD ab. Auch die Altersklasse der bis 30-Jährigen ist in unserer UG und KG deutlich stärker besetzt, als dies bei einer repräsentativen Abbildung der Bevölkerungsstruktur zu erwarten gewesen wäre. Die Gruppe der 40- bis 50-Jährigen ist hingegen leicht, die Gruppe der über 50-jährigen deutlich unterrepräsentiert.

b) Prüfung der Strukturgleichheit der Gruppen hinsichtlich der geschlechtlichen Zusammensetzung:

**Tabelle 7.2: Geschlechtliche Zusammensetzung von Untersuchungsgruppe und Kontrollgruppe, erhoben zum Messzeitpunkt t1**

	UG		KG	
	N	Prozent	N	Prozent
<b>M</b>	42	29,0	14	27,5
<b>W</b>	103	71,0	37	72,5
<b>Gesamt</b>	145	100,0	51	100,0

Tabelle 7.2 zeigt die geschlechtliche Zusammensetzung von Untersuchungs- und Kontrollgruppe. Es zeigt sich eine große Übereinstimmung beider Gruppen. Frauen sind in beiden Gruppen mit über 70 % vertreten, in der Untersuchungsgruppe ist der Anteil der Männer mit 29 % etwas höher als in der Kontrollgruppe. Tschuschke u. Anbeh (2000) berichten von der sogenannten ‚Zwei-Drittel-Quote‘, die besagt, dass ca. zwei Drittel der sich in psychotherapeutische Behandlung begebenden Patienten üblicherweise weiblichen Geschlechts sind. Die von uns vorgefundenen Häufigkeiten weichen von dieser Aussage nur unwesentlich ab.

c) Prüfung der Strukturgleichheit der Gruppen hinsichtlich des Schulabschlusses der Gruppenmitglieder:

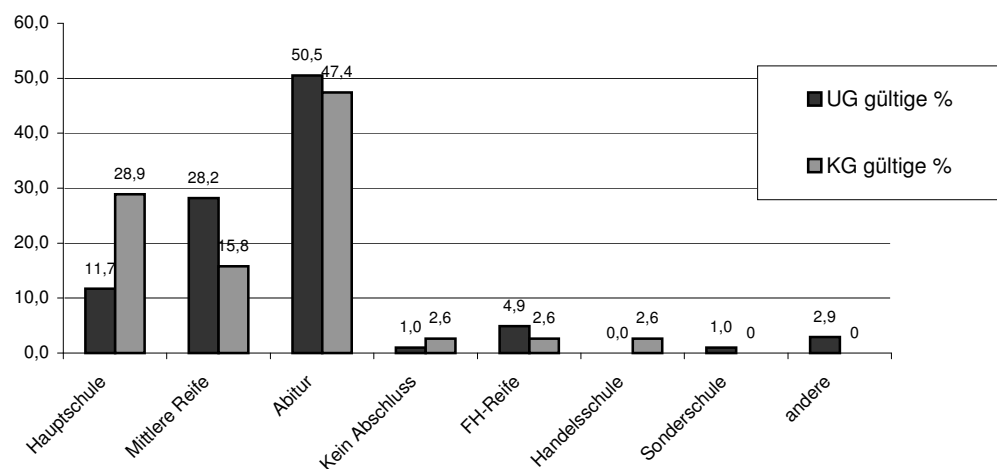
**Tabelle 7.3: Zusammensetzung von Untersuchungsgruppe und Kontrollgruppe nach Schulabschluss, erhoben zum Messzeitpunkt t1**

	UG			KG		
	N	%	gültige %	N	%	gültige %
Hauptschule	12	9,0	11,7	11	22,0	28,9
Mittlere Reife	29	21,6	28,2	6	12,0	15,8
Abitur	52	38,8	50,5	18	36,0	47,4
Kein Abschluss	1	0,7	1,0	1	2,0	2,6
FH-Reife	5	3,7	4,9	1	2,0	2,6
Handelsschule	- -	- -	- -	1	2,0	2,6
Sonderschule	1	0,7	1,0	- -	- -	- -
andere	3	2,2	2,9	- -	- -	- -
Gesamt	103	76,9	100,0	38	76,0	100,0
Fehlend	31	23,1		12	24,0	
Gesamt	134	100,0		50	100,0	

Tabelle 7.3 zeigt: Bei Zugrundelegung des Schulabschlusses der Patienten ergibt sich für die Untersuchungsgruppe mit 50 % wie auch für die Kontrollgruppe mit 47 % eine eindeutige Schwerpunktbildung in der Gruppe mit ‚Abitur‘.

In der Untersuchungsgruppe steht der Abschluss ‚Mittlere Reife‘ mit fast 30 % an zweiter Stelle, in der Kontrollgruppe ist der ‚Hauptschulabschluss‘ der zweithäufigste Schulabschluss mit ebenfalls fast 30 %. Demnach ist das Niveau der Schulbildung in der Untersuchungsgruppe etwas höher als in der Kontrollgruppe. Eine graphische Darstellung der Tabelle 7.3 zeigt Abbildung 10:

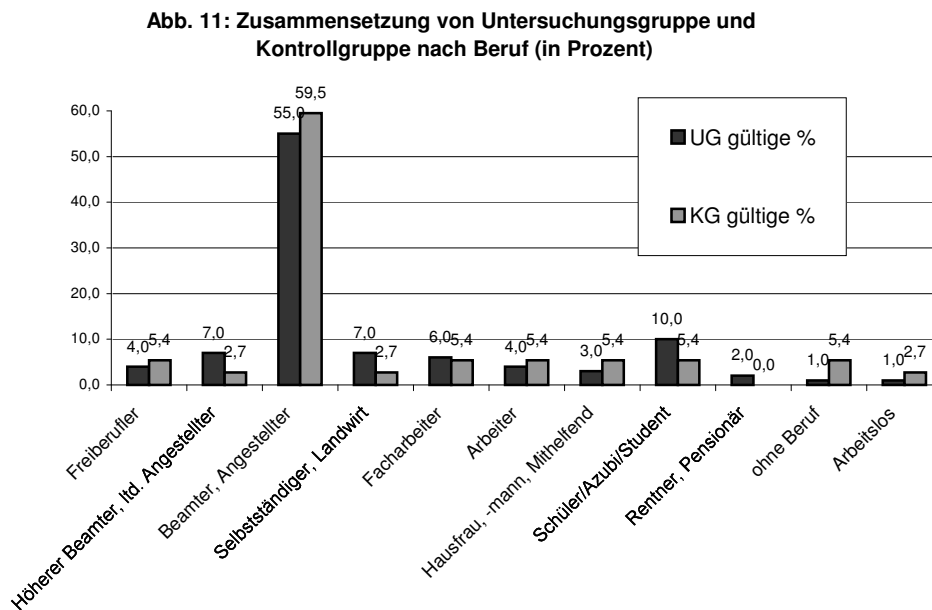
**Abb. 10: Zusammensetzung Untersuchungsgruppe und Kontrollgruppe nach Schulabschluss (in Prozent)**



d) Prüfung der Strukturgleichheit der Gruppen hinsichtlich des Berufs der Gruppenmitglieder:

Tabelle 7.4 zeigt hinsichtlich des Berufs der Patienten eine recht große Übereinstimmung zwischen Untersuchungs- und Kontrollgruppe. Über die Hälfte der Patienten gehört zur Berufsgruppe ‚Beamter, Angestellter‘, wobei der Anteil dieser Berufsgruppe in der Kontrollgruppe etwas höher liegt als in der Untersuchungsgruppe.

Die Gruppen ‚Schüler/Azubi/Student‘ und die Gruppe ‚Höherer Beamter, leitender Angestellter‘ und ‚Selbstständiger, Landwirt‘ sind dagegen in der Untersuchungsgruppe etwas stärker vertreten als in der Kontrollgruppe. Eine graphische Darstellung der Tabelle 7.4 zeigt Abbildung 11:



**Tabelle 7.4: Zusammensetzung von Untersuchungsgruppe und Kontrollgruppe nach Beruf der Patienten, erhoben zum Messzeitpunkt t1**

	UG			KG		
	N	%	gültige %	N	%	gültige %
Freiberufler	4	3,0	4,0	2	4,0	5,4
Höherer Beamter, ltd. Angestellter	7	5,2	7,0	1	2,0	2,7
Beamter, Angestellter	55	41,0	55,0	22	44,0	59,5
Selbstständiger, Landwirt	7	5,2	7,0	1	2,0	2,7
Facharbeiter	6	4,5	6,0	2	4,0	5,4
Arbeiter	4	3,0	4,0	2	4,0	5,4
Hausfrau, -mann, Mithelfend	3	2,2	3,0	2	4,0	5,4
Schüler/Azubi/Student	10	7,5	10,0	2	4,0	5,4
Rentner, Pensionär	2	1,5	2,0	- -	- -	- -
ohne Beruf	1	0,7	1,0	2	4,0	5,4
Arbeitslos	1	0,7	1,0	1	2,0	2,7
Gesamt	100	74,6	100,0	37	74,0	100,0
Fehlend	34	25,4		13	26,0	
Gesamt	134	100,0		50	100,0	

Zur Höhe der „N“ in den Tabellen 7.3 u.7.4 ist anzumerken: In den beiden Kategorien ‚Schulabschluss‘ und ‚Beruf‘ ist die Zahl fehlender Daten relativ groß. Dies kommt dadurch zustande, dass diese Daten nach Ablauf der Datenerhebungsphase nacherhoben werden mussten. Versehentlich und unbemerkt

sind die entsprechenden Fragen im Untersuchungsmaterial der Erhebungsphase nämlich nicht enthalten gewesen.

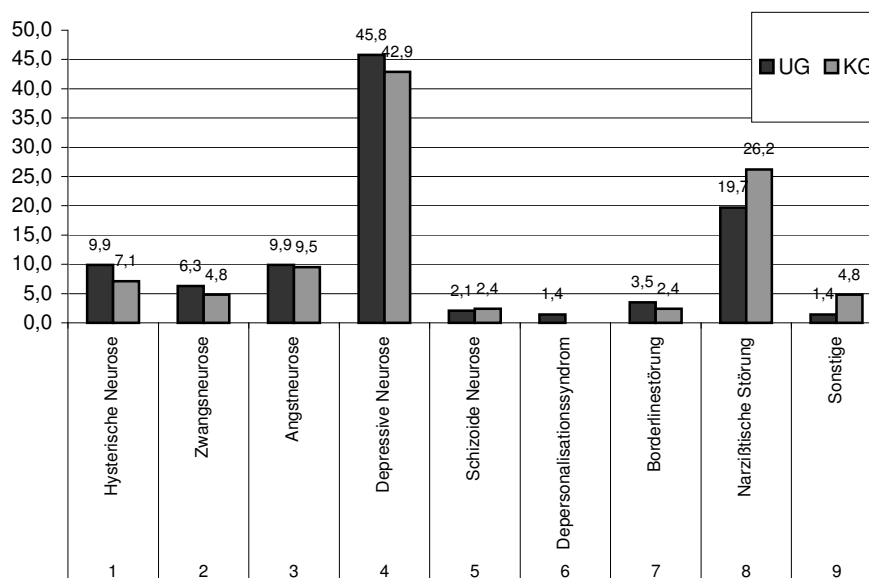
e) Prüfung der Strukturgleichheit der Gruppen hinsichtlich der in ihnen vertretenen Krankheitsbilder:

**Tabelle 7.5: Zusammensetzung von Untersuchungsgruppe und Kontrollgruppe nach Diagnosen, erhoben zum Messzeitpunkt t1**

	Diagnose	Häufigkeit		Prozent		Gültige Prozepte		Kumulierte Prozepte	
		UG	KG	UG	KG	UG	KG	UG	KG
<b>Gültig</b>	<b>Hysterische Neurose</b>	14	3	9,7	5,9	9,9	7,1	9,9	7,1
	<b>Zwangsneurose</b>	9	2	6,2	3,9	6,3	4,8	16,2	11,9
	<b>Angstneurose</b>	14	4	9,7	7,8	9,9	9,5	26,1	21,4
	<b>Depressive Neurose</b>	65	18	44,8	35,3	45,8	42,9	71,8	64,3
	<b>Schizoide Neurose</b>	3	1	2,1	2,0	2,1	2,4	73,9	66,7
	<b>Depersonalisationssyndrom</b>	2		1,4		1,4		75,4	
	<b>Borderlinestörung</b>	5	1	3,4	2,0	3,5	2,4	78,9	69,0
	<b>Narzißtische Störung</b>	28	11	19,3	21,6	19,7	26,2	98,6	95,2
	<b>Sonstige</b>	2	2	1,4	3,9	1,4	4,8	100,0	100,0
	<b>Gesamt</b>	142	42	97,9	82,4	100,0	100,0		
<b>Fehlend</b>		3	9	2,1	17,6				
<b>Gesamt</b>		145	51	100,0	100,0				

Tabelle 7.5 macht deutlich: bei Zugrundelegung der Diagnosen, so wie sie von den mitwirkenden Therapeuten angegeben wurden, stellt sich sowohl in der Untersuchungs- wie auch in der Kontrollgruppe eine Schwerpunktbildung bei der Diagnose vier (depressive Neurose) heraus, während die anderen Diagnosen, - außer Diagnose acht (narzißtische Persönlichkeitsstörung) -, nur selten vertreten sind. Diagnose acht ist in der Kontrollgruppe prozentual stärker vertreten als in der Untersuchungsgruppe. Eine graphische Darstellung der Tabelle 7.5 zeigt Abbildung 12:

**Abb. 12: Zusammensetzung von Untersuchungsgruppe und Kontrollgruppe nach Diagnosen (in Prozent)**



f) Prüfung der Strukturgleichheit der Gruppen hinsichtlich der in ihnen vertretenen Diagnosegruppen:

Wir haben unsere neun Diagnosestellungen in die internationale Klassifikation psychischer Störungen (ICD-10, 2. Auflage) überführt und erhalten dann die folgenden Gruppen: Erstens die große Gruppe der „affektiven Störungen“ (F 3), zu der auch die depressive Neurose (F 34.1) zählt; zweitens die große Gruppe der „neurotischen, belastungs- und somatoformen Störungen“ (F 4), in die wir die hysterischen Konversionsneurosen (F 44), Zwangsneurosen (F 42), Angstneurosen (F 41) und das Depersonalisationssyndrom (F 48.1) eingeordnet haben; drittens die große Gruppe der „Persönlichkeits- und Verhaltensstörungen“ (F 6) mit den „spezifischen Persönlichkeitsstörungen“ (F 60) und unseren Diagnosen „Schizoide Neurose“ (F 60.1), „Borderlinestörung“ (F 60.3) und „Narzisstische Persönlichkeitsstörung“ (F 60.8); viertens: unsere Rubrik „Sonstige Störungen“ wurde keiner der ICD-Gruppen zugeordnet.

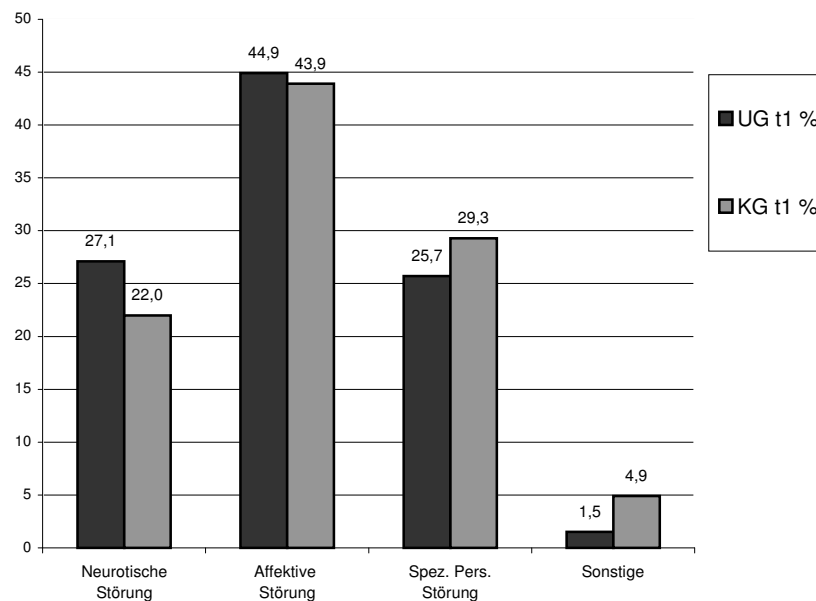
**Tabelle 7.6: Zusammensetzung von Untersuchungsgruppe und Kontrollgruppe nach Diagnosegruppen des ICD-10, erhoben zum Messzeitpunkt t1**

	UG		KG	
	t1		t1	
	N	%	N	%
<b>Neurotische Störung</b>	38	27,1	9	22,0
<b>Affektive Störung</b>	61	44,9	18	43,9
<b>Spez. Pers. Störung</b>	35	25,7	12	29,3
<b>Sonstige</b>	2	1,5	2	4,9
<b>Gesamt</b>	136	100	41	100,0

Tabelle 7.6 zeigt: In der Untersuchungsgruppe sind die affektiven Störungen (depressive Neurose) am stärksten vertreten, - dagegen machen die neurotischen Störungen und die „spezifischen Persönlichkeitsstörungen“ ungefähr nur je 1/4 der Diagnosen aus. In der Kontrollgruppe sind die Diagnosegruppen ähnlich verteilt, - mit einem leicht höheren Anteil der „spezifischen Persönlichkeitsstörungen“ und einen entsprechend geringeren Anteil der „neurotischen Störungen“. Eine graphische Darstellung der Tabelle 7.6 zeigt Abbildung 13:



**Abb. 13: Zusammensetzung von Untersuchungsgruppe und Kontrollgruppe nach Diagnosegruppen (in Prozent)**



g) Prüfung der Strukturgleichheit der Gruppen hinsichtlich der Zugehörigkeit der Therapeuten zu bestimmten Ausbildungsinstituten: An welchem Institut haben die Therapeuten von Untersuchungsgruppe und Kontrollgruppe ihre Ausbildung gemacht?

**Tabelle 7.7: Zusammensetzung der Untersuchungsgruppe und der Kontrollgruppe nach dem Ort des Ausbildungsinstituts des/der Therapeuten/in**

Ort	UG	KG
Aachen	25,5	29,4
Düsseldorf	31,7	25,5
Delmenhorst	18,6	15,7
München	21,4	17,6
unbest.	2,8	11,8

Tabelle 7.7 zeigt: Die vier Ausbildungsinstitute der DGIP in Aachen, Düsseldorf, Delmenhorst und München sind unterschiedlich stark vertreten, wobei sich die Patienten der Untersuchungsgruppe mit einer leichten Schwerpunktbildung mit über 30 % bei Therapeuten finden, die ihre Ausbildung in Düsseldorf gemacht haben. Die Kontrollgruppenpatienten kommen zu 29 % von Therapeuten, die ihre Ausbildung in Aachen gemacht haben. Ansonsten aber ist die prozentuale Beteiligung der Ausbildungsinstitute an der Untersuchungs- und Kontrollgruppe relativ ausgewogen.

Möglicherweise spielt bei der institutsbezogen unterschiedlichen Mitwirkung der Therapeuten eine Rolle, dass die Mitglieder der Forschungsgruppe aus dem Köln/Düsseldorfer Raum stammen und selbst Mitglieder des Düsseldorfer Instituts sind.

Zusammenfassung des Kapitels: Unsere Ergebnisse zeigen, dass wir zwar nicht von einer idealtypischen, so doch von einer relativ großen Strukturgleichheit (Homogenität) von Untersuchungs- und Kontrollgruppe ausgehen können; unter Vernachlässigung kleinerer Unterschiede gibt es in den untersuchten Kriterien zumindest keine gravierenden Abweichungen zwischen den beiden Gruppen. Somit möchten wir die Voraussetzung für eine Vergleichbarkeit beider Gruppen als gegeben ansehen. Eventuelle Messunterschiede in den abhängigen Variablen werden deshalb von uns nicht auf strukturelle Gruppenunterschiede in den hier untersuchten Kriterien zurückgeführt.

Nachdem nun die Voraussetzungen für einen Mittelwertsvergleich geklärt sind, können die t-Tests durchgeführt werden.

#### **10.3.2.2 t-Tests mit den Mittelwertunterschieden von Untersuchungsgruppe und Kontrollgruppe zu den Messzeitpunkten t1, t2 und t3 (Querschnittuntersuchungen)**

Zur Überprüfung der Annahme, dass die Untersuchungsgruppe nach einem halben und nach einem ganzen Jahr psychotherapeutischer Behandlung in den abhängigen Variablen statistisch abgesichert günstigere Ergebnisse als die Kontrollgruppe zeigt, wurden t-Tests durchgeführt. Zur Bestimmung des Störungsniveaus vor der Behandlung wurden ebenfalls t-Tests gerechnet, wobei wir die Vermutung hatten, dass sich Untersuchungs- und Kontrollgruppe hinsichtlich des Störungsgrades zu diesem Zeitpunkt noch nicht voneinander unterscheiden würden, - was ja auch eine Voraussetzung für die Rückführung späterer Veränderungen auf die intervenierende Therapievariable ist.

**Tabelle 9: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede von Untersuchungsgruppe und Kontrollgruppe zum Messzeitpunkt t1 (Querschnittvergleich) unter Berücksichtigung des Levene-Tests zur Prüfung der Varianzgleichheit von zwei unabhängigen Stichproben**

						Levene-Test der Varianzgleichheit			
		N	Mittelwert	StA*		F	Signifikanz	Sig. (2-seitig)	Mittlere Differenz
<b>GBB1-Erschöpfungsneigung</b>	UG	144	10,63	5,86	Varianzen sind gleich	0,002	0,966	0,225	1,1650
	KG	50	9,46	5,76	Varianzen sind nicht gleich			0,223	1,1650
<b>GBB2-Magenbeschwerden</b>	UG	143	5,37	4,44	Varianzen sind gleich	1,326	0,251	0,634	0,3314
	KG	51	5,04	3,70	Varianzen sind nicht gleich			0,604	0,3314
<b>GBB3-Gliederschmerzen</b>	UG	141	8,61	5,40	Varianzen sind gleich	0,019	0,892	0,680	0,3705
	KG	51	8,98	5,71	Varianzen sind nicht gleich			0,688	0,3705
<b>GBB4-Herzbeschwerden</b>	UG	141	5,88	4,00	Varianzen sind gleich	0,738	0,391	0,576	0,3755
	KG	51	6,25	4,39	Varianzen sind nicht gleich			0,594	0,3755
<b>GBB5-Beschwerdedruck</b>	UG	134	30,03	14,88	Varianzen sind gleich	0,731	0,394	0,842	0,5099
	KG	50	29,52	16,61	Varianzen sind nicht gleich			0,849	0,5099
<b>MMPI-Lie</b>	UG	142	3,35	1,96	Varianzen sind gleich	0,491	0,484	0,212	0,3949
	KG	50	3,74	1,80	Varianzen sind nicht gleich			0,196	0,3949
<b>MMPI-Validity</b>	UG	139	5,01	3,87	Varianzen sind gleich	0,081	0,776	0,337	0,6294
	KG	45	4,38	3,61	Varianzen sind nicht gleich			0,322	0,6294
<b>MMPI-Correction</b>	UG	140	8,51	3,36	Varianzen sind gleich	0,121	0,728	0,709	0,2183
	KG	45	8,29	3,55	Varianzen sind nicht gleich			0,717	0,2183
<b>MMPI-Hypochondriasis</b>	UG	141	9,43	4,42	Varianzen sind gleich	0,065	0,798	0,909	0,0900
	KG	50	9,52	5,18	Varianzen sind nicht gleich			0,908	0,0900
<b>MMPI-Depression</b>	UG	144	13,00	3,39	Varianzen sind gleich	5,095	0,025	0,184	1,0780
	KG	49	15,58	3,28	Varianzen sind nicht gleich			0,232	1,0278
<b>MMPI-Hysteria</b>	UG	137	15,58	3,94	Varianzen sind gleich	0,257	0,613	0,891	0,1256
	KG	48	15,46	4,01	Varianzen sind nicht gleich			0,895	0,1256
<b>MMPI-Psychopathie</b>	UG	136	12,74	4,42	Varianzen sind gleich	3,445	0,065	0,018	1,9496
	KG	42	10,79	5,18	Varianzen sind nicht gleich			0,031	1,9496
<b>MMPI-maskulin-feminin</b>	UG	133	15,01	3,39	Varianzen sind gleich	0,103	0,748	0,188	0,7422
	KG	49	14,27	3,28	Varianzen sind nicht gleich			0,184	0,7422
<b>MMPI-Paranoia</b>	UG	134	8,18	3,94	Varianzen sind gleich	0,273	0,602	0,184	0,8874
	KG	48	7,29	4,01	Varianzen sind nicht gleich			0,190	0,8874
<b>MMPI-Psychasthenia</b>	UG	140	14,77	5,11	Varianzen sind gleich	5,220	0,023	0,054	1,7306
	KG	49	13,04	6,08	Varianzen sind nicht gleich			0,078	1,7306
<b>MMPI-Schizoidie</b>	UG	133	9,42	4,87	Varianzen sind gleich	1,350	0,247	0,224	1,0381
	KG	47	8,38	5,41	Varianzen sind nicht gleich			0,250	1,0381
<b>MMPI-Hypomanie</b>	UG	134	10,4	3,16	Varianzen sind gleich	1,269	0,261	0,394	0,4883
	KG	46	10,89	3,83	Varianzen sind nicht gleich			0,439	0,4883
<b>MMPI-Soziale Introversion</b>	UG	141	13,84	4,6	Varianzen sind gleich	7,500	0,007	0,231	0,9797
	KG	49	12,86	5,73	Varianzen sind nicht gleich			0,283	0,9797

\* StA = Standardabweichung

Die Tabelle 9 macht die weitgehende Strukturgleichheit hinsichtlich des Störungsniveaus zum Messzeitpunkt t1, also für die UG vor der Behandlung und für die KG zu Beginn der Wartezeit, - gemessen mit den fünf Skalen des GBB und den dreizehn klinischen Skalen des MMPI(K) -, deutlich. Lediglich in zwei Skalen, in MMPI(K)-PP und -PT, liegen statistisch bedeutsame Unterschiede zwischen UG und KG vor. Auf Grund dieser weitgehenden ‚Störungsgleichheit‘ zu Beginn der Behandlung kann davon ausgegangen werden, dass sich die zu späteren Messzeitpunkten gemessenen Unterschiede zwischen UG und KG nicht mit einem bereits zu Beginn der Untersuchung bzw. vor der Behandlung

vorhandenen statistisch bedeutsamen unterschiedlichen Ausgangsniveau an Morbidität beider Gruppen erklären lassen (s. auch Tabelle 7.5 u. 7.6).

**Tabelle 10: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede von Untersuchungsgruppe und Kontrollgruppe zum Messzeitpunkt t2 (Querschnittvergleich) unter Berücksichtigung des Levene-Tests zur Prüfung der Varianzgleichheit von zwei unabhängigen Stichproben**

						Levene-Test der Varianzgleichheit		Sig. (2-seitig)	Mittlere Differenz
						F	Signifikanz		
GBB1-Erschöpfungsneigung	UG	124	8,77	5,74	Varianzen sind gleich	0,003	0,953	0,838	0,21
	KG	41	8,98	5,51	Varianzen sind nicht gleich			0,835	0,21
GBB2-Magenbeschwerden	UG	124	3,94	3,56	Varianzen sind gleich	5,136	0,025	0,136	-1,02
	KG	41	4,95	4,35	Varianzen sind nicht gleich			0,181	-1,02
GBB3-Gliederschmerzen	UG	122	6,89	4,58	Varianzen sind gleich	2,976	0,086	0,026	-1,96
	KG	41	8,85	5,53	Varianzen sind nicht gleich			0,045	-1,96
GBB4-Herzbeschwerden	UG	124	4,59	3,65	Varianzen sind gleich	6,832	0,010	0,137	-1,07
	KG	41	5,66	4,84	Varianzen sind nicht gleich			0,200	-1,07
GBB5-Beschwerdedruck	UG	122	24,10	13,26	Varianzen sind gleich	6,726	0,010	0,096	-4,34
	KG	41	28,00	17,35	Varianzen sind nicht gleich			0,148	-4,34
MMPI-Lie	UG	116	3,57	1,73	Varianzen sind gleich	0,652	0,421	0,734	-0,11
	KG	40	3,68	1,61	Varianzen sind nicht gleich			0,725	-0,11
MMPI-Validity	UG	117	4,39	2,92	Varianzen sind gleich	0,104	0,747	0,652	-0,26
	KG	32	4,66	2,91	Varianzen sind nicht gleich			0,653	-0,26
MMPI-Correction	UG	121	9,08	3,65	Varianzen sind gleich	0,024	0,877	0,692	0,28
	KG	35	8,80	3,90	Varianzen sind nicht gleich			0,703	0,28
MMPI-Hypochondriasis	UG	123	7,89	4,39	Varianzen sind gleich	2,254	0,135	0,503	-0,57
	KG	39	8,46	5,21	Varianzen sind nicht gleich			0,541	-0,57
MMPI-Depression	UG	122	12,56	4,70	Varianzen sind gleich	1,484	0,225	0,645	0,42
	KG	37	12,14	5,41	Varianzen sind nicht gleich			0,670	0,42
MMPI-Hysteria	UG	121	13,61	4,93	Varianzen sind gleich	1,501	0,222	0,691	-0,39
	KG	37	14,00	5,95	Varianzen sind nicht gleich			0,720	-0,39
MMPI-Psychopathie	UG	118	12,06	4,52	Varianzen sind gleich	3,135	0,079	0,138	1,33
	KG	37	10,73	5,36	Varianzen sind nicht gleich			0,178	1,33
MMPI-maskulin-feminin	UG	121	14,62	3,54	Varianzen sind gleich	0,286	0,593	0,139	0,95
	KG	39	13,67	3,29	Varianzen sind nicht gleich			0,127	0,95
MMPI-Paranoia	UG	120	7,20	3,51	Varianzen sind gleich	3,391	0,067	0,752	0,23
	KG	36	6,97	4,62	Varianzen sind nicht gleich			0,786	0,23
MMPI-Psychasthenia	UG	121	13,16	5,26	Varianzen sind gleich	5,069	0,026	0,443	0,78
	KG	40	12,38	6,48	Varianzen sind nicht gleich			0,492	0,78
MMPI-Schizoidie	UG	117	8,30	4,60	Varianzen sind gleich	0,046	0,830	0,895	-0,12
	KG	36	8,42	4,88	Varianzen sind nicht gleich			0,899	-0,12
MMPI-Hypomanie	UG	121	10,19	3,46	Varianzen sind gleich	0,448	0,504	0,887	0,10
	KG	33	10,09	3,89	Varianzen sind nicht gleich			0,895	0,10
MMPI-Soziale Introversion	UG	121	14,18	4,67	Varianzen sind gleich	0,331	0,566	0,553	0,54
	KG	36	13,64	5,25	Varianzen sind nicht gleich			0,579	0,54
VEV	UG	124	185,16	28,31	Varianzen sind gleich	0,944	0,333	0,011	13,60
	KG	41	171,56	32,46	Varianzen sind nicht gleich			0,020	13,60

Tabelle 10 macht deutlich: Im VEV zeigen sich zwischen t1 und t2, also nach einem halben Jahr der ‚Behandlung‘ bzw. des ‚Wartens‘ signifikante Unterschiede zwischen UG und KG in der Einschätzung des Therapieerfolgs, - was jedoch auch nicht anders zu erwarten war, denn die KG hat zu diesem Zeitpunkt ja auch noch keine Therapieerfahrungen, die sie bewerten könnte.

Die UG zeigt nach einem halben Jahr bereits eine behandlungsbedingte positivere Selbsteinschätzung hin zu mehr ‚Gelassenheit, Entspannung und Optimismus im Erleben und Verhalten‘ als zu Beginn der Behandlung zum Messzeitpunkt t1.

Im GBB zeigen im Vergleichszeitraum t1 zu t2 besonders zwei Skalen positiv-abweichende Entwicklungen der UG im Vergleich zur KG an: in der Skala „Gliederschmerzen“ ist der Unterschied zur KG „signifikant“, in der Skala „Beschwerdedruck“ ist er „tendenziell signifikant“. In allen anderen Skalen, auch in allen MMPI/K-Skalen, zeigen sich dagegen keine statistisch bedeutsamen Unterschiede im direkten Vergleich zwischen beiden Gruppen, d.h., dass sich lediglich geringe auf die Behandlung zurückführbare Veränderungen zeigen. Offensichtlich sind in der UG nach einem halben Jahr Behandlungszeit nur marginale Abweichungen von der KG feststellbar. Ganz anders verhält sich dies nach einem ganzen Jahr der Behandlung, wie Tabelle 11 (vgl. S. 107) deutlich macht: Zu t3 weisen alle GBB-Skalen in der UG mindestens signifikant verbesserte Mittelwerte im Vergleich zur Kontrollgruppe auf. Offensichtlich hat die psychotherapeutische Behandlung deutliche Besserungen in den psychosomatischen Symptom-Skalen hervorgebracht. Auch im VEV ist der Unterschied zwischen der Untersuchungsgruppe und der unbehandelten Kontrollgruppe „sehr signifikant“, d.h. es zeigt sich in der UG im Gegensatz zur KG ein großer Zugewinn an ‚Gelassenheit, Entspannung und Optimismus im Erleben und Verhalten‘, der mit hoher Wahrscheinlichkeit auf die psychotherapeutische Behandlung zurückzuführen ist, weil die morbide Ausgangslage beider Gruppen zum Messzeitpunkt t1 ja in etwa gleich war und außer der „Therapiewirkung“ keine anderen intervenierenden Variablen bekannt sind. Es zeigt sich, dass die in der KG durch die wenigen Kontakte zwischen Therapeut und Patient bewirkten Veränderungen nicht ausreichend sind, um mit der therapiebedingten Entwicklung der UG mithalten zu können.

Ein nicht ganz so günstiges Bild bekommen wir durch die Messwerte der Skalen des MMPI(K) vermittelt: die Untersuchungsgruppe weist zwar durchweg niedrigere Mittelwerte als die Kontrollgruppe auf, - ein Signifikanzniveau erreichen diese Unterschiede zwischen UG und KG jedoch nur in den Skalen Hypochondriasis und Schizoidie. Es drängen sich verschiedene Erklärungen auf: Die erste Möglichkeit betrifft die Qualität des Messinstruments: der MMPI(K) zeigt die stattgefundenen Veränderungen nicht so deutlich an wie der GBB, weil der GBB ein sensibleres Messinstrument ist. Die zweite Möglichkeit betrifft den Zeitverzug der Bewusstwerdung beim Patienten: therapeutisch hervorgerufene Veränderungen bilden sich in körpernahen psychosomatischen Beschwerdeskalen früher als in körperfernen psychischen Beschwerdeskalen ab, weil die Linderung von körperlichen Beschwerden den Patienten schneller auffällt als innerpsychische Veränderungen. Die dritte Möglichkeit lässt die Ergebnisse uninterpretiert so wie sie sind: körperliche Beschwerden verändern sich durch Therapie eher als psychische Beschwerden.

Unabhängig von möglichen Erklärungen bleibt das statistische Faktum festzuhalten, dass die oben formulierte Annahme hinsichtlich der vermuteten Überlegenheit psychotherapeutischer Behandlung im Vergleich zur vermuteten Unwirksamkeit eines bloßen ‚Wartens‘ im Vergleichszeitraum t1 zu t3 zumindest mit dem MMPI(K) nur in zwei von 13 Skalen bestätigt werden konnte.

**Tabelle 11: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede von Untersuchungsgruppe und Kontrollgruppe zum Messzeitpunkt t3 (Querschnittvergleich) unter Berücksichtigung des Levene-Tests zur Prüfung der Varianzgleichheit von zwei unabhängigen Stichproben**

						Levene-Test der Varianzgleichheit		Sig. (2-seitig)	Mittlere Differenz
						F	Signifikanz		
GBB1-Erschöpfungsneigung	UG	110	7,35	5,26	Varianzen sind gleich	5,116	0,025	0,039	-2,20
	KG	38	9,55	6,47	Varianzen sind nicht gleich			0,064	-2,20
GBB2-Magenbeschwerden	UG	110	3,55	3,40	Varianzen sind gleich	2,511	0,115	0,011	-1,75
	KG	37	5,30	4,13	Varianzen sind nicht gleich			0,024	-1,75
GBB3-Gliederschmerzen	UG	110	5,96	4,14	Varianzen sind gleich	7,470	0,007	0,005	-2,51
	KG	38	8,47	6,03	Varianzen sind nicht gleich			0,021	-2,51
GBB4-Herzbeschwerden	UG	109	3,83	3,80	Varianzen sind gleich	8,460	0,004	0,048	-1,61
	KG	38	5,45	5,51	Varianzen sind nicht gleich			0,101	-1,61
GBB5-Beschwerdedruck	UG	106	20,58	12,40	Varianzen sind gleich	9,143	0,003	0,003	-8,37
	KG	37	28,95	18,65	Varianzen sind nicht gleich			0,014	-8,37
MMPI-Lie	UG	100	3,80	1,81	Varianzen sind gleich	0,085	0,771	0,407	0,29
	KG	37	3,51	1,73	Varianzen sind nicht gleich			0,398	0,29
MMPI-Validity	UG	94	4,31	3,29	Varianzen sind gleich	0,258	0,612	0,614	-0,35
	KG	32	4,66	3,58	Varianzen sind nicht gleich			0,630	-0,35
MMPI-Correction	UG	109	9,70	3,74	Varianzen sind gleich	0,513	0,475	0,128	1,09
	KG	36	8,61	3,56	Varianzen sind nicht gleich			0,122	1,09
MMPI-Hypochondriasis	UG	109	7,28	4,50	Varianzen sind gleich	1,987	0,161	0,045	-1,80
	KG	38	9,08	5,38	Varianzen sind nicht gleich			0,069	-1,80
MMPI-Depression	UG	108	10,89	5,12	Varianzen sind gleich	2,937	0,089	0,270	-1,14
	KG	37	12,03	6,13	Varianzen sind nicht gleich			0,315	-1,14
MMPI-Hysteria	UG	104	12,28	5,90	Varianzen sind gleich	1,271	0,261	0,122	-1,83
	KG	37	14,11	6,79	Varianzen sind nicht gleich			0,151	-1,83
MMPI-Psychopathie	UG	103	10,56	4,55	Varianzen sind gleich	4,869	0,029	0,845	-0,19
	KG	36	10,75	5,88	Varianzen sind nicht gleich			0,863	-0,19
MMPI-maskulin-feminin	UG	109	14,73	3,10	Varianzen sind gleich	1,984	0,161	0,980	1,06
	KG	37	13,68	3,98	Varianzen sind nicht gleich			0,147	1,06
MMPI-Paranoia	UG	104	6,78	3,65	Varianzen sind gleich	0,376	0,541	0,792	-0,19
	KG	36	6,97	4,13	Varianzen sind nicht gleich			0,804	-0,19
MMPI-Psychasthenia	UG	109	11,93	5,58	Varianzen sind gleich	2,021	0,157	0,810	-0,27
	KG	35	12,20	6,63	Varianzen sind nicht gleich			0,827	-0,27
MMPI-Schizoidie	UG	105	7,28	4,68	Varianzen sind gleich	0,078	0,781	0,070	-1,75
	KG	33	9,03	5,21	Varianzen sind nicht gleich			0,090	-1,75
MMPI-Hypomanie	UG	104	9,90	3,67	Varianzen sind gleich	2,185	0,142	0,277	-0,85
	KG	33	10,76	4,62	Varianzen sind nicht gleich			0,338	-0,85
MMPI-Soziale Introversion	UG	109	12,93	4,71	Varianzen sind gleich	3,820	0,053	0,528	0,60
	KG	37	12,32	5,78	Varianzen sind nicht gleich			0,569	0,60
VEV	UG	108	195,73	37,06	Varianzen sind gleich	0,005	0,943	0,002	22,55
	KG	39	173,18	38,83	Varianzen sind nicht gleich			0,003	22,55

Zusammenfassung: Ein Jahr nach Therapiebeginn lässt sich im Kontrollgruppenvergleich die Therapiewirksamkeit für die Untersuchungsgruppe in allen GBB-Skalen, im VEV, sowie in zwei von 13 betrachteten MMPI(K)-Skalen nachweisen. Im VEV und in zwei der fünf GBB-Skalen waren die erwarteten Unterschiede zugunsten der UG schon nach einem halben Jahr nachweisbar.

### **10.3.3 Effektivitätsfeststellung durch Unterschiedsmessungen diverser Prä-Post-Vergleiche (Längsschnittuntersuchungen) mit Hilfe von GBB-Skalen, MMPI(K)-Skalen sowie durch VEV-Vergleiche im dreijährigen Untersuchungszeitraum**

Im folgenden Schritt möchten wir nicht mehr UG und KG miteinander (Querschnittsuntersuchung), sondern wir möchten die Entwicklung der beiden Gruppen in sich, also die Veränderung innerhalb der jeweiligen beiden Gruppen durch Vergleich ihrer klinischen Symptomausprägungen zwischen den aufeinander folgenden Messzeitpunkten (Längsschnittuntersuchung) untersuchen.

Um dies sinnvoll durchführen zu können, muss wieder eine Voraussetzung gegeben sein, nämlich dass sich die Gruppenstruktur im Laufe der Zeit nicht so verändert, dass eventuelle Messergebnisse nicht mehr auf die Therapiewirkung, sondern u.U. auf die Veränderung der Gruppenstruktur zurückzuführen sind. Beispiel: Wenn vornehmlich die besonders kranken Patienten im Laufe der Untersuchungszeit, - aus welchen Gründen auch immer -, aus der Untersuchung ausscheiden, also nur noch die vergleichsweise „gesunden Kranken“ von der Untersuchung erfasst werden, dann drückt eine u.U. gemessene „Leidensminderung“ nicht mehr aus, dass die Therapie gewirkt hat, sondern die „Leidensminderung“ kann dann lediglich Ausdruck des Umstandes sein, dass zu dem betreffenden Messzeitpunkt nur noch die von Anfang an bereits weniger symptombelasteten Patienten, die sich im Extremfall überhaupt nicht verändert haben müssten, Auskünfte geben und das Messergebnis so gesehen natürlich günstiger ausfällt, als wenn auch die Antworten der schwerer symptombelasteten Patienten hinzukämen.

Es ist also zu klären, ob sich die Gruppenstruktur von Untersuchungs- bzw. Kontrollgruppe durch das Ausscheiden von Patienten im Laufe der Untersuchungszeit wesentlich verändert hat oder nicht. Nur wenn die Gruppenstruktur über die Messzeitpunkte t1 bis t5 in etwa gleich geblieben ist, können eventuelle gemessene klinische Unterschiede auf die intervenierende Variable „Psychotherapie“ zurückgeführt werden. Hat sich jedoch die Gruppenstruktur verändert, dann können die gemessenen klinischen Veränderungen auch Ausdruck der Strukturveränderung der Gruppe sein.

Wir prüfen die strukturelle Homogenität der Untersuchungs- wie der Kontrollgruppe anhand des zentralen Merkmals „Diagnosenstruktur“ und fragen, ob die Anteile der Diagnosen zu den verschiedenen Testzeitpunkten gleich geblieben sind. Wenn dies bei im Kriterium „Diagnosevergleich“ zutrifft, unterstellen wir ohne Prüfung anderer Merkmale Strukturgleichheit der Gruppen und führen eventuelle Mittelwertunterschiede in den klinischen Skalen der Messinstrumente auf die intervenierende Variable (= psychotherapeutische Behandlung) zurück.

Darüber hinaus ist die Frage nach der über die Zeit stabilen vs. veränderten Diagnosestruktur auch im Hinblick auf die Prüfung der Frage nach einer diagnosespezifischen Indikation zur Langzeitpsychotherapie von Bedeutung, - meint konkret: Die Beobachtung der Veränderung der Diagnosestruktur im Zeitverlauf kann Hinweise auf die Beantwortung der Frage geben, welche Krankheitsgruppen kürzer, welche länger in Behandlung sind, bzw. welche Krankheitsgruppen eine kürzere und welche eine längere Behandlung benötigen.

### 10.3.3.1 Prüfung der diagnostischen Strukturgleichheit von Untersuchungsgruppe (UG) und Kontrollgruppe (KG) als Voraussetzung der Aussagegültigkeit von Prä-Post-Vergleichen (Längsschnittvergleichen)

Wir führen den Strukturvergleich der Diagnoseverteilung in mehreren Schritten durch:

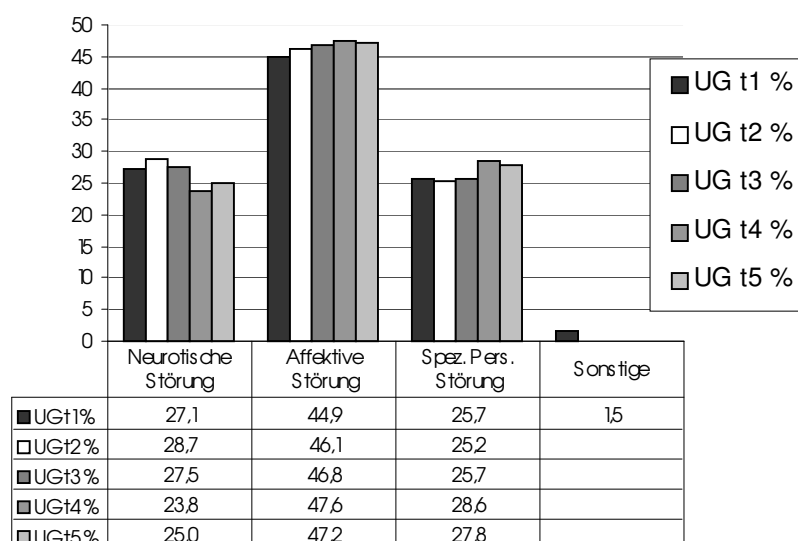
a) Zunächst haben wir den Anteil der *Diagnosegruppen* (nach ICD-10) in der Untersuchungsgruppe zu den verschiedenen Testzeitpunkten miteinander verglichen.

**Tabelle 12.1: Zusammensetzung der Untersuchungsgruppe nach Diagnosegruppen im Zeitverlauf von t1 bis t5**

	UG									
	t1		t2		t3		t4		t5	
	N	%	N	%	N	%	N	%	N	%
Neurotische Störung	38	27,1	33	28,7	30	27,5	20	23,8	18	25,0
Affektive Störung	61	44,9	53	46,1	51	46,8	40	47,6	34	47,2
Spez. Pers. Störung	35	25,7	29	25,2	28	25,7	24	28,6	20	27,8
Sonstige	2	1,5								
Gesamt	136	100	115	100	109	100	84	100	72	100

Tabelle 12.1 zeigt die Anteile der Diagnosegruppen (maßgebliche Diagnosestellung durch die behandelnden Therapeuten jeweils zum Messzeitpunkt t1) in der Untersuchungsgruppe (UG). Eine graphische Darstellung zeigt Abbildung 14:

**Abb. 14: Zusammensetzung der Untersuchungsgruppe nach Diagnosegruppen im Zeitverlauf von Messzeitpunkt t1 bis Messzeitpunkt t5 (in Prozent)**





Die Diagnosegruppen bleiben über die fünf Messzeitpunkte gesehen prozentual relativ konstant, tendenziell zeigt sich jedoch eine prozentuale Abnahme der „neurotischen Störungen“ und eine Zunahme der „affektiven Störungen“ (depressive Neurose) sowie der „spezifischen Persönlichkeitsstörungen“ bis zum Zeitpunkt t5. Dies ist vielleicht ein Hinweis darauf, dass die beiden letztgenannten Störungsgruppen tendenziell länger behandelt werden als die erstgenannte Störungsgruppe, vielleicht auch ein Hinweis darauf, dass die beiden letztgenannten Störungsgruppen eher zu den „schwereren Störungen“ zu zählen sind.

b) Auch bezüglich der Kontrollgruppe stellt sich die Frage nach der strukturellen Konstanz der diagnostischen Verteilung, und zwar sowohl für den Zeitraum während der Wartezeit, also bis zum Messzeitpunkt t3, wie auch für die darauf folgenden Behandlungszeiträume (t3 bis t5), in denen die Kontrollgruppenpatienten automatisch zu Untersuchungsgruppenpatienten werden.

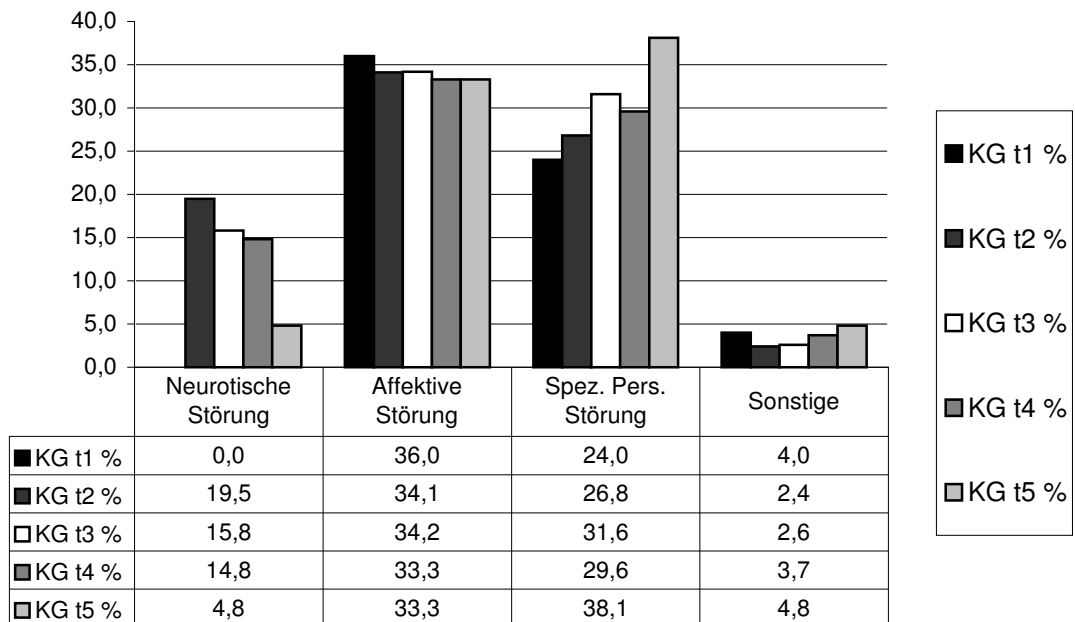
**Tabelle 12.2: Zusammensetzung der Kontrollgruppe nach Diagnosegruppen im Zeitverlauf von t1 bis t5**

	KG									
	t1		t2		t3		t4		t5	
	N	%	N	%	N	%	N	%	N	%
<b>Neurotische Störung</b>	9	18,0	8	19,5	6	15,8	4	14,8	1	4,8
<b>Affektive Störung</b>	18	36,0	14	34,1	13	34,2	9	33,3	7	33,3
<b>Spez. Pers. Störung</b>	12	24,0	11	26,8	12	31,6	8	29,6	8	38,1
<b>Sonstige</b>	2	4,0	1	2,4	1	2,6	1	3,7	1	4,8
<b>Gesamt</b>	41	82,0	34	82,9	32	84,2	22	81,5	17	81,0
<b>Fehlend</b>	9	18,0	7	17,1	6	15,8	5	18,5	4	19,0
<b>Gesamt</b>	50	100,0	41	100,0	38	100,0	27	100,0	21	100,0

Tabelle 12.2 zeigt: Die Anteile der Diagnosegruppen (Diagnosestellung jeweils zu t1) in der Kontrollgruppe (KG) entwickeln sich bis zum Zeitpunkt t3 (Ende der Wartezeit) zwar unterschiedlich, dennoch werden die prozentualen Größenordnungen der Verteilung der Diagnosegruppen hiervon nicht wesentlich berührt. Der Anteil der „depressiven Neurose“ bleibt konstant, während der Anteil der „neurotischen Störung“ etwas abnimmt und die „spezifische Persönlichkeitsstörung“ prozentual etwas zunimmt. Diese Entwicklung setzt sich auch in der nach der „Wartezeit“ folgenden zweijährigen Behandlungszeit (t4 und t5) fort. Auch hier zeigt sich wieder, dass die schwereren Störungen - auch schon in der Wartezeit - eher an der Behandlung (bzw. der Behandlungserwartung) festhalten, während die weniger schweren neurotischen Störungen – aus welchem Grund auch immer – die Behandlung früher beenden.

Eine graphische Darstellung zeigt Abbildung 15:

**Abb. 15: Zusammensetzung der Kontrollgruppe nach Diagnosegruppen im Zeitverlauf von Messzeitpunkt t1 bis Messzeitpunkt t5 (in Prozent)**



c) Zur Beantwortung der Frage nach der Strukturgleichheit der Untersuchungsgruppe im Zeitverlauf von t1 bis t5 sowie der Frage, ob Störungsgruppen *länger oder kürzer in Behandlung* bleiben, haben wir im ersten Schritt die diagnostischen Einschätzungen der Therapeuten als Kriterium herangezogen und die so gebildeten statistischen Diagnosegruppen in der Entwicklung ihrer prozentualen Proportionen zueinander im Zeitverlauf von t1 bis t5 verglichen.

Nunmehr betrachten wir empirisch die von diesen diagnostischen Gruppen in den klinischen Skalen erzielten Mittelwerte zu den verschiedenen Messzeitpunkten. Wir greifen hierzu beispielhaft auf die fünf Skalen des GBB und die vier zentralen Skalen des MMPI(K) (drei Skalen der neurotischen Trias zzgl. PT) sowie die Messzeitpunkte t3 und t5 heraus. Die Daten der folgenden Tabelle basieren dabei für t3 auf der statistischen Subgruppe „vollständige Daten pro Skala sowohl für t1 wie auch für t3 vorliegend“ sowie für t5 auf der statistischen Subgruppe „vollständige Daten pro Skala sowohl für t3 wie auch für t5 vorliegend“.

**Tabelle 13: Vergleich der Mittelwerte der drei Diagnosegruppen\* in den GBB-Skalen und den MMPI(K)-Skalen HS, D, HY und PT auf der Basis der Grundgesamtheit zu t3 und zu t5**

		t1-t3		t3-t5		Differenz
		N	M	N	M	(t1-t3)-(t3-t5)
GBB 1	N	30	7,233	15	8,400	-1,167
GBB 1	A	50	6,920	32	6,937	-0,017
GBB 1	P	28	8,071	18	8,278	-0,207
GBB 1	G	108	7,306	65	7,646	-0,340
GBB 2	N	31	3,581	15	4,133	-0,552
GBB 2	A	51	3,588	31	3,677	-0,089
GBB 2	P	26	3,346	17	2,824	0,522
GBB 2	G	108	3,528	63	3,556	-0,028
GBB 3	N	28	5,750	15	6,733	-0,983
GBB 3	A	49	5,918	32	6,531	-0,613
GBB 3	P	28	5,464	18	4,611	0,853
GBB 3	G	105	5,752	65	6,046	-0,294
GBB 4	N	29	4,931	15	6,000	-1,069
GBB 4	A	50	3,180	32	3,812	-0,632
GBB 4	P	26	3,385	18	2,941	0,444
GBB 4	G	105	3,714	65	4,094	-0,380
GBB 5	N	26	20,346	15	25,267	-4,921
GBB 5	A	47	19,404	31	21,161	-1,757
GBB 5	P	25	20,120	16	18,188	1,932
GBB 5	G	98	19,837	62	21,387	-1,550
HS	N	28	6,929	15	7,733	-0,804
HS	A	50	7,320	31	7,581	-0,261
HS	P	27	7,704	17	6,647	1,057
HS	G	109	7,314	63	7,365	-0,051
D	N	29	9,897	15	10,200	-0,303
D	A	50	10,620	31	11,226	-0,606
D	P	28	12,536	18	10,778	1,758
D	G	107	10,925	64	10,859	0,066
HY	N	29	11,138	15	12,067	-0,929
HY	A	43	11,837	29	12,241	-0,404
HY	P	27	14,370	17	12,353	2,017
HY	G	99	12,323	61	12,230	0,093
PT	N	31	11,258	15	11,600	-0,342
PT	A	47	11,660	31	12,581	-0,921
PT	P	26	13,423	17	12,000	1,423
PT	G	104	12,113	63	12,060	0,053
* N= Neurotische Störung		P= Spezifische Persönlichkeitsstörung				
A= Affektive Störung		G= Gesamt				

Die Abnahme der „N“ von t3 zu t5 basiert auf Therapiebeendigungen planmäßiger oder unplanmäßiger Art. Die Frage ist, ob durch das Ausscheiden von Patienten der Mittelwert der verbleibenden Patienten in den drei Diagnosegruppen größer oder kleiner wird.

Die Datenanalyse ergibt folgendes Bild: bei den neurotischen Störungen steigen die Mittelwerte durch das Ausscheiden von Patienten in der Größenordnung von ca. 0,3 – 4,9 Punkten, bei den depressiven

Störungen von ca. 0,02 – 1,7 Punkten an, wohingegen die Werte in der Diagnosegruppe „Persönlichkeitsstörung“ in der Größenordnung von ca. 0,4 – 2,0 Punkten (mit einer Ausnahme eines Anstiegs in der GBB-1-Skala/Erschöpfungsneigung von 0,207) fallen. Dies kann bedeuten, dass in der Gruppe der „persönlichkeitsgestörten Patienten“ die schwerer gestörten Patienten die Behandlung früher beendet (u.U. abgebrochen) haben als die leichter gestörten Patienten, während in der Gruppe mit der Diagnose „neurotische“ bzw. „affektiv-depressive Störung“ die leichter gestörten Patienten die Behandlung früher beendet haben, - vielleicht auch nicht mehr behandlungsbedürftig waren -, so dass die Mittelwerte der verbleibenden Patienten zwangsläufig ansteigen.

d) Im nächsten Schritt vergleichen wir die Gruppe der „ausgeschiedenen Patienten“ mit der Gruppe der „verbliebenen Patienten“. Frage: Unterscheiden sich die Patienten, die vor dem Zeitpunkt t5 ausgeschieden sind (Gruppe A) von denen, die bis zu diesem Zeitpunkt in der Untersuchungsgruppe geblieben sind (Gruppe B) hinsichtlich ihrer Diagnosen? Zur Beantwortung dieser Frage haben wir beide Gruppen hinsichtlich ihrer Diagnosen (Diagnosegruppen) miteinander verglichen. Die beiden Gruppen können statistisch durch das Merkmal voneinander getrennt werden, ob zum Messzeitpunkt t5 noch vollständige Datensätze im GBB und im MMPI(K) vorliegen oder nicht. Liegen noch Daten vor, heißt dies, dass es sich um Patienten handelt, welche die Therapie zu diesem Messzeitpunkt noch nicht beendet haben, - liegen keine Daten mehr vor, haben diese Patienten die Therapie zu diesem Messzeitpunkt beendet. Die Erfassung der Patienten kann statistisch über eine beliebige Skala der beiden Testverfahren erfolgen, - wir wählten die GBB-Skala 5 Beschwerdedruck und die MMPI(K)-Skala Depression.

**Tabelle 14: Vergleich der ausgeschiedenen Patientengruppe (A) mit der bis zum Messzeitpunkt t5 der Untersuchung verbliebenen Patientengruppe (B) hinsichtlich ihrer zum Messzeitpunkt t1 festgestellten Diagnosen (Diagnosegruppe)**

	Neurotische Störung		Affektive Störung		spezifische Persönlichkeitsstörung		Gesamt	
	N	%	N	%	N	%	N	%
<b>A</b>	23	32,4	32	45,1	16	12,5	71	100,0
<b>B</b>	16	23,2	33	47,8	20	29,0	69	100,0
<b>Gesamt</b>	39		65		36		140	

Tabelle 14 zeigt: Es ist festzuhalten, dass beide Gruppen hinsichtlich der prozentualen Größenordnungen grundsätzlich Strukturgleichheit zeigen. Auf der deskriptiven Ebene kann man jedoch bei genauerem Hinsehen einen Trend feststellen: Es zeigt sich, dass der Anteil der „affektiven Störungen“ in beiden Gruppen ungefähr gleich ist, - allerdings mit einem leichten Übergewicht in der Gruppe der „verbliebenen Patienten“, „neurotische Störungen“ sind in der Gruppe der „ausgeschiedenen Patienten“, „spezifische Persönlichkeitsstörungen“ in der Gruppe der „verbliebenen Patienten“ stärker vertreten. Dies könnte ein Hinweis darauf sein, dass Patienten mit der Diagnose einer „spezifischen Persönlichkeitsstörung“ länger in Behandlung bleiben als die Patienten mit „neurotischer Störung“. Diese sind in der Gruppe der Patienten, die zum Messzeitpunkt t5 keine Daten mehr vorliegen haben, stärker vertreten, d. h. wahrscheinlich kürzer in Behandlung gewesen.

Dies widerspricht dem Befund von Schulz et al. (1999), die bei Patienten einer stationären psychosomatischen Rehabilitation feststellen konnten, dass in der Diagnosegruppe „Persönlichkeitsstörungen“ die Rate frühzeitiger Entlassungen 1,5 – 2 mal höher lag als bei anderen Patienten.

*Zusammenfassend* kann festgestellt werden, dass trotz der zu den verschiedenen Messzeitpunkten leichten Unterschiede in der Diagnosestruktur der Untersuchungsgruppe als auch der Kontrollgruppe letztlich keine so gravierenden Verschiebungen der Häufigkeitsverteilungen der Diagnosen festgestellt werden konnten, als dass von einer wirklich bedeutsamen Strukturveränderung hinsichtlich ihrer diagnostischen Zusammensetzung gesprochen werden könnte. Allerdings gibt es gewisse – nicht widerspruchsfreie - Tendenzen, dass nämlich Patienten mit der Diagnose „neurotische Störung“ eher kürzer, hingegen Patienten mit einer „affektiven Störung“ (depressive Neurose) oder „spezifischen Persönlichkeitsstörung“ tendenziell eher länger in Behandlung bleiben.

Nachdem die Voraussetzungen für die inhaltliche Interpretierbarkeit der Längsschnittvergleiche geklärt sind, können nunmehr sinnvoll t-Tests für abhängige Stichproben zur Prüfung der Mittelwertunterschiede innerhalb der Untersuchungs- und der Kontrollgruppe vorgenommen werden. In der Untersuchungsgruppe (UG) wurden t-Tests zu den Messzeitpunktvergleichen t1:t2, t2:t3, t3:t4, t4:t5 und t3:t5, in der Kontrollgruppe (KG) für die Messzeitpunktvergleiche t1:t2, t2:t3 und t1:t3 durchgeführt.

#### **10.3.3.2 t-Tests zur Signifikanzprüfung der Mittelwertunterschiede der fünf Messzeitpunkte in der Untersuchungsgruppe (diverse Prä-Post-Längsschnittvergleiche)**

Zur Prüfung der Annahme, dass sich in der Untersuchungsgruppe mit zunehmender Dauer der Behandlung zwischen zwei Messzeitpunkten statistisch bedeutsame Veränderungen in den abhängigen Variablen zeigen, wurden t-Tests mit den Mittelwertunterschieden der verschiedenen Testzeitpunkte durchgeführt:

**Tabelle 15: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede zwischen Messzeitpunkt t1 und Messzeitpunkt t2, gemessen in den fünf GBB-Skalen, den Skalen HS, D, HY, PT des MMPI(K) der Untersuchungsgruppe**

		N	M	s	MD (t1/t2)	s (t1/t2)	Signif.	Signiveau*
<b>GBB1-Eerschöpfungsneigung</b>	<b>t1</b>	123	10,683	5,999	1,951	4,823	0,000	h.s.
	<b>t2</b>	123	8,730	5,750				
<b>GBB2-Magenbeschwerden</b>	<b>t1</b>	123	5,228	4,304	1,268	2,634	0,000	h.s.
	<b>t2</b>	123	3,960	3,560				
<b>GBB3-Gliederschmerzen</b>	<b>t1</b>	118	8,509	5,422	1,831	4,028	0,000	h.s.
	<b>t2</b>	118	6,680	4,440				
<b>GBB4-Herzbeschwerden</b>	<b>t1</b>	120	5,733	3,957	1,192	3,399	0,000	h.s.
	<b>t2</b>	120	4,540	3,620				
<b>GBB5-Beschwerdedruck</b>	<b>t1</b>	113	29,655	15,046	6,027	10,987	0,000	h.s.
	<b>t2</b>	113	23,630	13,150				
<b>MMPI-Hypochondriasis</b>	<b>t1</b>	120	9,267	4,752	1,350	3,127	0,000	h.s.
	<b>t2</b>	120	7,920	4,440				
<b>MMPI-Depression</b>	<b>t1</b>	121	14,116	4,416	1,554	3,396	0,000	h.s.
	<b>t2</b>	121	12,560	4,720				
<b>MMPI-Hysteria</b>	<b>t1</b>	116	15,474	5,284	1,902	3,840	0,000	h.s.
	<b>t2</b>	116	13,570	5,020				
<b>MMPI-Psychasthenia</b>	<b>t1</b>	117	14,675	5,184	1,479	3,323	0,000	h.s.
	<b>t2</b>	117	13,200	5,220				

Die Tabelle 15 zeigt die Entwicklung der Mittelwerte von t1 nach t2 mit den dazugehörigen Standardabweichungen in der Untersuchungsgruppe, gemessen in den fünf GBB-Skalen und den vier zentralen Skalen des MMPI(K) (drei Skalen der neurotische Trias zzgl. PT). Die Signifikanzprüfungen der Mittelwertunterschiede mit Hilfe des t-Tests zeigen, dass die Abnahme der Mittelwerte in allen überprüften Skalen ein „hochsignifikantes“ Niveau erreicht. Das heißt, dass schon im ersten halben Jahr bedeutsame Veränderungen in der Untersuchungsgruppe stattgefunden haben, und zwar sowohl was die Körpersymptome als auch was die neurotischen Störungen anbelangt.

Nach Rudolf (1994) handelt es sich bei symptomwertigen Skalen (z.B. allgemeine Ängstlichkeit, Herzklopfen, innere Anspannung, Schwindelgefühle, Atemnot, Kloßgefühl, Nackenschmerzen, Kopfschmerzen usw.) in besonderer Weise um veränderungssensible Messinstrumente.

Unsere Ergebnisse stimmen mit Ergebnissen von Kächele (1990) überein, der zeigen konnte, dass schon in den ersten psychotherapeutischen Sitzungen erhebliche Besserungen erreicht werden können.

**Tabelle 16: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede zwischen Messzeitpunkt t2 und Messzeitpunkt t3, gemessen in den fünf GBB-Skalen, den Skalen HS, D, HY, PT des MMPI(K) und im VEV der Untersuchungsgruppe**

		N	M	s	MD (t2/t3)	s (t2/t3)	Signif.	Signiveau*
GBB1-Erschöpfungsneigung	t2	100	9,12	5,56	1,83	4,14	0,000	h.s.
	t3	100	7,29	5,25				
GBB2-Magenbeschwerden	t2	100	3,85	3,44	0,53	2,27	0,021	s.
	t3	100	3,32	3,29				
GBB3-Gliederschmerzen	t2	99	6,89	4,50	0,93	3,12	0,004	s.s.
	t3	99	5,96	4,29				
GBB4-Herzbeschwerden	t2	99	4,49	3,30	0,76	2,69	0,006	s.s.
	t3	99	3,74	3,86				
GBB5-Beschwerdedruck	t2	95	24,31	12,62	4,12	8,27	0,000	h.s.
	t3	95	20,19	12,76				
MMPI-Hypochondriasis	t2	98	7,85	4,26	0,54	2,93	0,071	t.s.
	t3	98	7,31	4,65				
MMPI-Depression	t2	96	12,59	4,64	1,61	3,03	0,000	h.s.
	t3	96	10,98	5,34				
MMPI-Hysteria	t2	91	13,54	4,78	0,99	3,55	0,009	s.s.
	t3	91	12,55	6,03				
MMPI-Psychasthenia	t2	96	13,17	5,02	1,22	3,55	0,001	s.s.
	t3	96	11,95	5,63				
VEV	t2	98	183,81	26,57	-12,27	31,40	0,000	s.s.
	t3	98	196,07	37,34				

\*s. = signifikant, s.s. = sehr signifikant, h.s. = hoch signifikant, t.s. = tendenziell signifikant

Tabelle 16 zeigt die Entwicklung von t2 bis t3, also im zweiten Halbjahr der Behandlung. Die Mittelwerte entwickeln sich weiterhin degressiv, jedoch nicht mehr wie im ersten Halbjahr durchgehend „hoch signifikant“, - eine „hoch signifikante“ Abnahme der Mittelwerte von t2 nach t3 zeigt sich nur noch bei GBB-Erschöpfung, GBB-Beschwerdedruck und MMPI(K)-Depression. Immerhin noch „sehr signifikant“ sind die Symptomminderungen in den Skalen GBB-Gliederschmerzen, GBB-Herzbeschwerden, MMPI(K)-Hysterie und MMPI(K)-Psychasthenie. Von einfacher „Signifikanz“ der Symptomminderung kann noch bei „GBB-Magenbeschwerden“ und von „tendenziell signifikant“ noch bei MMPI(K)-Hypochondriasis gesprochen werden.

Im VEV ist der Unterschied zwischen t2 und t3 „sehr signifikant“, was auf eine positive Veränderung in Richtung hin zu mehr ‚Gelassenheit, Entspannung und Optimismus im Erleben und Verhalten‘ der Patienten verweist.

Unter den Körpersymptomen sind im zweiten Behandlungshalbjahr die „Magenbeschwerden“ am veränderungsresistentesten, während „Erschöpfungsgrad“ und „Beschwerdedruck“ am deutlichsten abnehmen. Das heißt, dass die Patienten in diesem Zeitraum der Therapie wieder vermehrt gesunden, was dazu passt, dass in diesen Zeitraum auch die Abnahme von Depressionen, gemessen in der MMPI(K)-D-Skala, als auch die im VEV gemessene Veränderung des Erlebens und Verhaltens fällt.

**Tabelle 17: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede zwischen Messzeitpunkt t3 und Messzeitpunkt t4, gemessen in den fünf GBB-Skalen, den Skalen HS, D, HY, PT des MMPI(K) und im VEV der Untersuchungsgruppe**

		N	M	s	MD (t3/t4)	s (t3/t4)	Signif.	Signiveau*
GBB1-Erschöpfungsneigung	t3	79	7,18	5,12	0,54	4,43	0,279	n.s.
	t4	79	6,63	5,03				
GBB2-Magenbeschwerden	t3	76	3,00	3,02	0,36	2,74	0,261	n.s.
	t4	76	2,64	2,99				
GBB3-Gliederschmerzen	t3	78	5,56	4,21	-0,08	2,64	0,798	n.s.
	t4	78	5,64	4,52				
GBB4-Herzbeschwerden	t3	76	3,57	3,79	0,13	2,76	0,679	n.s.
	t4	76	3,43	3,44				
GBB5-Beschwerdedruck	t3	75	18,83	11,63	0,77	9,18	0,468	n.s.
	t4	75	18,05	12,28				
MMPI-Hypochondriasis	t3	78	7,27	4,49	1,14	3,14	0,002	s.s.
	t4	78	6,13	4,29				
MMPI-Depression	t3	75	10,61	5,09	0,35	3,29	0,364	n.s.
	t4	75	10,27	4,73				
MMPI-Hysteria	t3	71	12,10	5,89	1,27	3,89	0,008	s.s.
	t4	71	10,83	5,29				
MMPI-Psychasthenia	t3	74	11,72	5,77	0,70	3,83	0,119	n.s.
	t4	74	11,01	5,68				
VEV	t3	77	196,58	37,25	-9,70	31,29	0,008	s.s.
	t4	77	206,29	32,98				

\* n.s. = nicht signifikant, s.s. = sehr signifikant

Tabelle 17 macht deutlich: Der Einjahresvergleich von t3 nach t4 (also der Zeitraum des zweiten Behandlungsjahres) zeigt nur noch wenige signifikante Entwicklungen in den untersuchten abhängigen Variablen. Zwar finden sich in den Skalen MMPI(K)-Hysterie, MMPI(K)-Hypochondriasis sowie im VEV noch „sehr signifikante“ Mittelwertveränderungen, - in den anderen Skalen (außer in GBB-Gliederschmerzen) gehen die Mittelwerte jedoch nur noch so wenig zurück, dass diese Veränderungen kein Signifikanzniveau mehr erreichen.

Die Entwicklung der Patienten, so wie sie sich in der HS-Skala des MMPI(K) darstellt, kann ein Hinweis darauf sein, dass Patienten mit hypochondrischen Störungen längere Zeit für eine Veränderung brauchen: Hatten sie zwischen t2 und t3 nur eine schwach signifikante Veränderung aufzuweisen, holen sie zwischen den Messzeitpunkten t3 und t4 gewissermaßen auf.



Auch die Entwicklung in der Hysterie-Skala macht einen seit Beginn der Behandlung in Gang gekommenen nachhaltig anhaltenden weiteren Rückgang der konversionshysterischen Symptome deutlich.

**Tabelle 18: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede zwischen Messzeitpunkt t4 und Messzeitpunkt t5, gemessen in den fünf GBB-Skalen, den Skalen HS, D, HY, PT des MMPI(K) und im VEV der Untersuchungsgruppe**

		N	M	s	MD (t4/t5)	s (t4/t5)	Signif.	Signiveau*
GBB1-Erschöpfungsneigung	t4	64	6,95	5,31	0,09	3,89	0,848	n.s.
	t5	64	6,86	4,93				
GBB2-Magenbeschwerden	t4	60	3,25	3,47	0,22	2,90	0,565	n.s.
	t5	60	3,03	3,39				
GBB3-Gliederschmerzen	t4	63	6,11	4,80	0,11	3,58	0,806	n.s.
	t5	63	6,00	4,44				
GBB4-Herzbeschwerden	t4	62	3,85	3,87	0,77	2,78	0,032	s.
	t5	62	3,08	3,36				
GBB5-Beschwerdedruck	t4	60	19,83	13,50	0,87	8,28	0,421	n.s.
	t5	60	18,97	12,78				
MMPI-Hypochondriasis	t4	64	6,47	4,68	0,28	3,68	0,543	n.s.
	t5	64	6,19	4,54				
MMPI-Depression	t4	62	10,40	4,86	0,97	3,88	0,054	t.s.
	t5	62	9,44	5,29				
MMPI-Hysteria	t4	60	11,28	5,40	0,90	4,37	0,116	n.s.
	t5	60	10,38	5,63				
MMPI-Psychasthenia	t4	62	11,56	5,80	1,53	4,04	0,004	s.s.
	t5	62	10,03	5,47				
VEV	t4	64	209,09	32,81	4,03	42,49	0,451	n.s.
	t5	64	205,05	44,35				

\* n.s. = nicht signifikant, t.s. = tendenziell signifikant, s. = signifikant, s.s. = sehr signifikant

Tabelle 18 zeigt, dass sich die Mittelwerte vom Messzeitpunkt t4 zum Messzeitpunkt t5 zwar erneut noch einmal etwas absenken, jedoch fällt diese Symptomminderung teilweise nur noch marginal aus. „Sehr signifikant“ ist der Unterschied zwischen t4 und t5 nur noch in der Skala MMPI(K)-Psychasthenie, „signifikant“ in GBB-Herzbeschwerden und „tendenziell signifikant“ noch in MMPI(K)-Depression. Alle anderen Unterschiede sind nicht signifikant. Dies gilt auch für die Mittelwerte im VEV, dessen Mittelwert sich zu t5 im Vergleich zu t4 sogar wieder etwas im Sinne einer Verschlechterung von ‚Gelassenheit, Entspannung und Optimismus im Erleben und Verhalten‘ entwickelt.

Betrachtet man den Verlauf der MMPI(K)-Skala Psychasthenie und GBB-Herzbeschwerden über den gesamten Therapieverlauf, dann ergibt sich das Bild, dass es erst in der zweiten Hälfte des ersten Behandlungsjahres zu einem deutlichen „Therapieerfolg“ bei diesen Störungsbildern kam, - ein Erfolg, der dann im zweiten Behandlungsjahr stagnierte und sich nun im dritten Behandlungsjahr wieder beschleunigt hat. Dies gilt in eingeschränktem Maße auch für die depressiven Patienten.

Man gewinnt den Eindruck, dass es bei den Patienten mit einer zwanghaft-ängstlichen Symptombelastung nach einer rückläufigen Phase nun (zum Behandlungsende) wieder zu einem größeren Schritt hin zu einem Abbau von zwanghaften Strukturen und zu einem Mehr an Unabhängigkeit und Autonomie ge-

kommen ist, was sich bei Patienten mit Herzbeschwerden, - entsprechend unseren Messergebnissen in der Skala GBB-5 -, sicherlich im Sinne einer Minderung ihrer Beschwerden ausgewirkt haben dürfte.

**Tabelle 19: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede zwischen Messzeitpunkt t3 und Messzeitpunkt t5, gemessen in den fünf GBB-Skalen, den Skalen HS, D, HY, PT des MMPI(K) und im VEV der Untersuchungsgruppe**

		N	M	s	MD (t3/t5)	s (t3/t5)	Signif.	Signiveau.*
GBB1-Erschöpfungsneigung	t3	65	7,65	5,34	0,57	3,93	0,248	n.s.
	t5	65	7,08	5,14				
GBB2-Magenbeschwerden	t3	63	3,56	3,21	0,54	2,68	0,114	n.s.
	t5	63	3,02	3,10				
GBB3-Gliederschmerzen	t3	65	6,05	4,31	0,25	3,70	0,593	n.s.
	t5	65	5,80	4,25				
GBB4-Herzbeschwerden	t3	64	4,09	4,10	1,14	3,19	0,006	s.s.
	t5	64	2,95	2,84				
GBB5-Beschwerdedruck	t3	62	21,39	13,21	2,55	9,30	0,035	s.
	t5	62	18,84	12,42				
MMPI-Hypochondriasis	t3	63	7,37	4,64	1,06	3,46	0,018	s.
	t5	63	6,30	4,34				
MMPI-Depression	t3	64	10,86	4,97	1,23	4,42	0,029	s.
	t5	64	9,63	5,15				
MMPI-Hysteria	t3	61	12,23	5,93	1,54	4,64	0,012	s.
	t5	61	10,69	5,41				
MMPI-Psychasthenia	t3	63	12,19	5,52	1,94	4,13	0,000	s.s.
	t5	63	10,25	5,35				
VEV	t3	62	196,81	34,63	-7,29	48,35	0,240	n.s.
	t5	62	204,10	43,26				

\* n.s. = nicht signifikant, s. = signifikant, s.s. = sehr signifikant

Gemäß Tabelle 19 verringern sich zwar im Zweijahreszeitraum von t3 nach t5 die Mittelwerte von GBB-Erschöpfung, GBB-Magenbeschwerden und GBB-Gliederschmerzen, die Unterschiede sind jedoch nicht signifikant. „Sehr signifikant“ verringert sich dagegen der Mittelwert von GBB-Herzbeschwerden und „signifikant“ der Mittelwert von GBB-Beschwerdedruck. Auch in den MMPI(K)-Skalen zeigen sich statistisch bedeutsame Unterschiede, und zwar „signifikante“ Unterschiede bei MMPI(K)-Hypochondriasis, MMPI(K)-Depression und MMPI(K)-Hysterie und „sehr signifikante“ Unterschiede bei MMPI(K)-Psychasthenie. Im VEV erhöht sich der Mittelwert zwar um mehr als sieben Punkte, bleibt jedoch unter der statistisch bedeutsamen Signifikanzgrenze.

Wir denken, damit zu einem interessanten Ergebnis gekommen zu sein: Wie wir gesehen haben, findet die Entwicklung bei MMPI(K)-Psychasthenie, MMPI(K)-Depression und GBB-Herzbeschwerden hauptsächlich im dritten Behandlungsjahr statt, in den anderen Skalen eher im zweiten Behandlungsjahr. Dieses Ergebnis spricht den Langzeitpsychotherapien einen ganz generellen therapeutischen Sinn zu, insofern nämlich damit (zumindest bei den von uns untersuchten Patienten) nachgewiesen ist, dass sich die durch Therapie erzeugten Veränderungen nicht - wie oft von Befürwortern von Kurzzeittherapien behauptet - nur zu Beginn der Therapien abspielen. Wir konnten nachweisen, dass sich bei den genannten Störungsbildern auch im dritten Behandlungsjahr noch bedeutsame therapeutische Erfolge zeigen, - Erfolge mit denen man bei vorschneller Betrachtung im zweiten Behandlungsjahr nicht mehr unbedingt hätte

rechnen müssen. Zum andern ist dieses Ergebnis aber auch deshalb interessant, weil sich darin eine diagnosespezifische Indikation für Langzeittherapie bei zwanghaften, herzneurotischen und - mit Einschränkung - auch depressiven Störungsbildern abzeichnen könnte.

Vergleicht man das erste Therapiejahr mit dem darauf folgenden therapeutischen Zweijahreszeitraum, so zeigt sich, dass der Gesundungsprozess in den Störungsformen der Depression, Hysterie, Hypochondriasis und Psychasthenie (Phobie/Zwangsverhalten) weit mehr vorangegangen ist als bei Störungsbildern mit körperlichen Symptomen, die bei Betrachtung dieses Langzeitraumes vergleichsweise nur geringe Verbesserungen zeigen.

Von einer ähnlichen Beobachtung berichten Kordy et al. (1988), die bei einer Gruppe psychosomatischer Patienten fanden, dass bei diesen mit zunehmender Behandlungsdauer von bis zu dreieinhalb Jahren eine immer geringer werdende Symptomverbesserung feststellbar war. Bassler et al. (1995 b) fanden für die stationäre Psychotherapie, dass eine längere Behandlungsdauer nicht automatisch zu einem besseren Behandlungserfolg führt. Dies steht im Widerspruch zu den bei Bachrach (1991) referierten Ergebnissen der Studien des New York Psychoanalytic Insitutes, in denen ein starker Zusammenhang zwischen der Länge der Behandlung und dem „outcome“ festgestellt werden konnte. Auch Tschuschke (1997) stellt in seiner Kritik an Grawe fest, dass in tiefenpsychologischen Behandlungen kurzfristig zwar auf allen Ebenen ein signifikanter Symptomrückgang und eine subjektive Befindlichkeitsverbesserung von nachweislich bis zu 70 % erreichbar ist, dass davon aber strukturelle Veränderungen mit tiefergehenden Konfliktlösungen relativ unberührt erscheinen, weshalb diese sowohl eine längere Behandlungszeit als auch eine höhere Behandlungsdosis (Dichte) benötigten.

#### **10.3.3.3 t-Tests zur Signifikanzprüfung der Mittelwertunterschiede der drei Messzeitpunkte in der Kontrollgruppe (diverse Prä-Post-Längsschnittvergleiche)**

Unsere bisher ermittelten Ergebnisse in der Untersuchungsgruppe haben nur dann einen Wert, wenn nachgewiesen werden kann, dass sich in der gleichen Zeit in der nichtbehandelten Kontrollgruppe keine derartigen Veränderungen (z.B. als Spontanheilungen) einstellten. Nur wenn dies nachgewiesen werden kann, lassen sich ceteris-paribus die Ergebnisse in der Untersuchungsgruppe auf „Therapieeinwirkung“ zurückführen.

Zur Überprüfung der Annahme, dass sich die Patienten der Kontrollgruppe während des Wartens in den abhängigen Variablen (klinische Skalen der Messinstrumente) statistisch nicht signifikant verändern, wurden mit Hilfe von t-Tests Signifikanzprüfungen der Unterschiede zwischen den jeweiligen Messzeitpunkten vorgenommen.

**Tabelle 20: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede zwischen Messzeitpunkt t1 und Messzeitpunkt t2, gemessen in den fünf GBB-Skalen und den Skalen HS, D, HY, PT des MMPI(K) der Kontrollgruppe**

	N	M	S	MD (t1/t2)	S (t1/t2)	Signif.	Signiveau*
<b>GBB1-Erschöpfungsneigung</b>	t1 40	9,725	5,311	0,825	3,608	0,156	n.s.
	t2 40	8,900	5,560				
<b>GBB2-Magenbeschwerden</b>	t1 41	4,902	3,737	-0,048	2,429	0,898	n.s.
	t2 41	4,950	4,350				
<b>GBB3-Gliederschmerzen</b>	t1 41	9,049	5,594	0,195	3,318	0,709	n.s.
	t2 41	8,850	5,530				
<b>GBB4-Herzbeschwerden</b>	t1 41	6,342	4,644	0,683	3,110	0,167	n.s.
	t2 41	5,660	4,840				
<b>GBB5-Beschwerdedruck</b>	t1 40	29,750	16,100	1,425	8,910	0,318	n.s.
	t2 40	28,330	17,500				
<b>MMPI-Hypochondriasis</b>	t1 38	9,316	4,332	0,711	2,629	0,104	n.s.
	t2 38	8,610	5,200				
<b>MMPI-Depression</b>	t1 36	12,667	5,026	0,333	2,683	0,461	n.s.
	t2 36	12,330	5,350				
<b>MMPI-Hysteria</b>	t1 34	15,235	5,309	1,324	3,574	0,038	s.
	t2 34	13,910	6,110				
<b>MMPI-Psychasthenia</b>	t1 38	12,632	5,875	-0,211	3,223	0,690	n.s.
	t2 38	12,840	6,280				

\* n.s. = nicht signifikant, s. = signifikant

Tabelle 20 zeigt die in den fünf GBB-Skalen sowie den vier zentralen Skalen des MMPI(K) (drei Skalen der neurotischen Trias zzgl. PT) bei der Kontrollgruppe zu t1 und t2 gemessenen Mittelwerte, Standardabweichungen, mittleren Differenzen, Signifikanzen und die daraus resultierenden Signifikanzniveaus. Es zeigen sich durchweg keine signifikanten Unterschiede zwischen t1 und t2, mit Ausnahme eines signifikanten Wertes in der MMPI(K)-Skala Hysterie, in der die Werte von t1 nach t2 auch ohne Behandlung signifikant abnehmen. Alle anderen, - teilweise auch durchaus abnehmenden Skalenwerte -, sind signifikanzstatistisch gesehen nicht von Belang.

Es bleibt die Frage, wieso es in der Hysterie-Skala zu einer signifikanten Symptomverbesserung gekommen ist. An dieser Stelle kann zum Verständnis ein Befund von Franz et al. (2000) zur „Spontanheilung“ angeführt werden: Franz et al. stellten nämlich in einer Jahre andauernden Langzeitbeobachtung am Beispiel einer Bevölkerungsstichprobe fest, dass 23,1 Prozent der klinischen Beeinträchtigten einen günstigen Spontanverlauf aufweisen. Es kann sein, dass sich diese Tendenz bei uns rein zufällig in der Hysterie-Skala abgebildet hat. Es kann aber auch sein, dass diese Abnahme in der Hysterie-Skala bereits ein hysterieeignes Krankheitsmerkmal umschreibt, dass nämlich das „Hysterische“ dazu neigt, sich zunächst geradezu übertrieben zur Schau zu stellen, sich zu exhibitionieren, zu exaltieren, sich selbst zu inflationieren, um sich dann jedoch bei nächster Gelegenheit genau so wieder zurück zu nehmen. Zumindest wäre damit erklärbar, wieso sich in dieser Skala bei uns zunächst eine Mittelwertsteigerung, dann jedoch eine Mittelwertminderung zeigt. So gesehen wäre die „Signifikanz“ der Mittelwertabnahme in der Hysterie-Skala kein Umstand, der den Wert der Kontrollgruppe mindern würde.

Wir können unsere Überlegung überprüfen, indem wir beobachten, wie sich die Hysterie-Skala im nächsten Beobachtungszeitraum von t2 zu t3 verhält. Stimmt unsere Vermutung, dann dürfte sich nämlich zwischen t2 und t3 keine weitere wesentliche Mittelwertsenkung mehr zeigen, weil diese ja bereits erfolgt sein müsste.

**Tabelle 21: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede zwischen Messzeitpunkt t2 und Messzeitpunkt t3, gemessen in den fünf GBB-Skalen, den Skalen HS, D, HY, PT des MMPI(K) und im VEV der Kontrollgruppe**

		N	M	S	MD (t2/t3)	S (t2/t3)	Signif.	Signiveau*
<b>GBB1-Erschöpfungsneigung</b>	<b>t2</b>	35	9,40	5,54	-0,77	4,44	0,311	n.s.
	<b>t3</b>	35	10,17	6,29				
<b>GBB2-Magenbeschwerden</b>	<b>t2</b>	34	5,29	4,44	-0,03	3,11	0,956	n.s.
	<b>t3</b>	34	5,32	4,18				
<b>GBB3-Gliederschmerzen</b>	<b>t2</b>	35	9,11	5,67	0,34	3,69	0,586	n.s.
	<b>t3</b>	35	8,77	6,10				
<b>GBB4-Herzbeschwerden</b>	<b>t2</b>	35	6,11	5,05	0,66	3,90	0,326	n.s.
	<b>t3</b>	35	5,46	5,61				
<b>GBB5-Beschwerdedruck</b>	<b>t2</b>	34	29,88	18,15	-0,06	11,75	0,977	n.s.
	<b>t3</b>	34	29,94	18,76				
<b>MMPI-Hypochondriasis</b>	<b>t2</b>	34	8,82	5,28	-0,56	3,79	0,396	n.s.
	<b>t3</b>	34	9,38	5,31				
<b>MMPI-Depression</b>	<b>t2</b>	32	12,41	5,37	-0,41	3,95	0,565	n.s.
	<b>t3</b>	32	12,81	5,95				
<b>MMPI-Hysteria</b>	<b>t2</b>	31	14,19	6,25	-0,84	4,71	0,330	n.s.
	<b>t3</b>	31	15,03	6,48				
<b>MMPI-Psychasthenia</b>	<b>t2</b>	33	12,33	6,56	0,03	3,74	0,963	n.s.
	<b>t3</b>	33	12,30	6,63				
<b>VEV</b>	<b>t2</b>	36	171,56	32,46	-1,62	38,83	0,782	n.s.
	<b>t3</b>	36	173,18	38,83				

\* n.s. = nicht signifikant

Auch zwischen den Messzeitpunkten t2 und t3 sollten idealtypischerweise in der Kontrollgruppe eigentlich keine statistisch bedeutsamen Veränderungen gemessen werden. In Tabelle 21 zeigt sich, dass die Unterschiede in den Mittelwerten der GBB-Skalen und den MMPI(K)-Skalen tatsächlich auch eher gering sind. Zwar sind die Mittelwerte für GBB-Erschöpfung und für die MMPI(K)-Skalen (mit Ausnahme der Psychasthenie-Skala) leicht angestiegen, dafür sind die Mittelwerte bei GBB-Gliederschmerzen und GBB-Herzbeschwerden jedoch leicht zurückgegangen. Alle Veränderungen bewegen sich im statistisch „nicht signifikanten“ Bereich, und auch der VEV-Wert verändert sich nur unerheblich. Was die oben diskutierte Veränderung in der Hysterie-Skala anbelangt, so zeigt sich im jetzt beobachteten Zeitraum keine signifikante Mittelwertsenkung mehr, was als ein Plausibilitätsbeleg für unsere oben dargelegten Vermutungen gelten kann.

In einem weiteren Schritt werfen wir nun die Messergebnisse von t1 zu t2 und von t2 zu t3 zusammen und überblicken den Gesamtzeitraum von t1 zu t3:

**Tabelle 22: t-Tests zur Signifikanzprüfung der Mittelwertunterschiede zwischen Messzeitpunkt t1 und Messzeitpunkt t3, gemessen in den fünf GBB-Skalen und den Skalen HS, D, HY, PT des MMPI(K) der Kontrollgruppe**

		N	M	S	MD (t1/t3)	S (t1/t3)	Signifikanz (2-seitig)	Signif.-Niveau*
<b>GBB1-Erschöpfungsneigung</b>	<b>t1</b>	37	10,30	5,54	0,6486	5,4631	0,475	n.s.
	<b>t3</b>	37	9,65	6,53				
<b>GBB2-Magenbeschwerden</b>	<b>t1</b>	37	5,35	3,71	0,0514	2,7883	0,907	n.s.
	<b>t3</b>	37	5,30	4,13				
<b>GBB3-Gliederschmerzen</b>	<b>t1</b>	38	9,66	5,85	1,1842	4,5313	0,116	n.s.
	<b>t3</b>	38	8,47	6,03				
<b>GBB4-Herzbeschwerden</b>	<b>t1</b>	38	6,81	4,64	1,3684	3,2418	0,013	s.
	<b>t3</b>	38	5,45	5,51				
<b>GBB5-Beschwerdedruck</b>	<b>t1</b>	37	32,00	16,55	3,0541	12,6578	0,151	n.s.
	<b>t3</b>	37	28,95	18,65				
<b>MMPI-Hypochondriasis</b>	<b>t1</b>	38	9,74	4,58	0,6579	3,3148	0,229	n.s.
	<b>t3</b>	38	9,08	5,38				
<b>MMPI-Depression</b>	<b>t1</b>	37	13,03	5,07	1,0000	3,9861	0,136	n.s.
	<b>t3</b>	37	12,03	6,13				
<b>MMPI-Hysteria</b>	<b>t1</b>	37	15,65	5,64	1,5405	4,2921	0,036	s.
	<b>t3</b>	37	14,11	6,79				
<b>MMPI-Psychasthenia</b>	<b>t1</b>	34	12,74	5,99	0,2941	3,7701	0,652	n.s.
	<b>t3</b>	34	12,44	6,57				

\* n.s. = nicht signifikant, s. = signifikant

Tab. 22 zeigt: Die Prüfung der Mittelwertunterschiede zwischen t1 und t3 ergibt für die GBB-Herzbeschwerden-Skala und die MMPI(K)-Hysterie-Skala wiederum eine signifikante Abnahme, während in den anderen Skalen die Unterschiede zwischen t1 und t3 statistisch nicht von Bedeutung sind.

Zur Abnahme in der Hysterie-Skala ist zu sagen, dass sich diese ausschließlich aus der Entwicklung zwischen t1 und t2 erklärt (vgl. oben) und sich, - einmal stattgefunden -, im Mittelwertvergleich t1 zu t3 natürlich noch einmal abbildet, - ohne damit jedoch ein weiteres neues Ereignis darzustellen.

Zur Erklärung des gegenläufigen Verlaufs der „Herzbeschwerden“ im Vergleich zu den anderen chronischen Schmerzsyndromen greifen wir auf Greif-Higer u.a. (1993) zurück, die Überlegungen angestellt haben, dass Patienten mit einer Herzanxietyneurose möglicherweise ihre bloße Entscheidung für eine Psychotherapie bereits auch dann beschwerdemindernd als einen Gewinn an Autonomie empfinden, wenn dieser Schritt zunächst erst einmal in einer Wartegruppe endet.

Zusammenfassung: Unsere Ergebnisse können die oben formulierte Annahme nicht vorhandener Symptomverbesserungen in der Kontrollgruppe nicht durchgängig bestätigen. Offensichtlich gibt es sowohl im Halbjahres- wie im Jahresvergleich einige, - wenngleich wenige -, so doch nicht zu leugnende Symptomverbesserungen, die nicht auf Psychotherapie zurückzuführen sind. Bringt man diesen Befund in der Kontrollgruppe allerdings in einen Vergleich zu den Veränderungen in der Untersuchungsgruppe, dann spricht die Dominanz der Symptomverbesserungen in der Untersuchungsgruppe jedoch eindeutig und zweifelsfrei für die überlegene Wirkung von Psychotherapie im Vergleich zur Spontanheilung in der Kontrollgruppe.

#### **10.3.4 Varianzanalyse zur Signifikanzprüfung des unterschiedlichen Verlaufs von UG und KG zwischen den Messzeitpunkten t1 und t3 (vgl. Tabellen 79 a und 79 b im Anhang)**

In Kap. 10.3.2 ff. führten wir zum Messzeitpunkt t1, t2 und t3 jeweils eine Signifikanzprüfung der Mittelwerte von UG und KG als Querschnittvergleich durch. In Kap. 10.3.3 ff. prüften wir die Signifikanz der Mittelwerte der fünf aufeinander folgenden Messzeitpunkte in der Untersuchungsgruppe sowie der drei aufeinander folgenden Messzeitpunkte in der Kontrollgruppe, - jeweils als Längsschnittvergleich.

Alle diese Untersuchungen wurden mit dem t-Test gerechnet. Dabei wurden immer zwei Datengruppen miteinander verglichen: UG/t1 zu KG/t1, UG/t2 zu KG/t2, UG/t3 zu KG/t3, UG/t1 zu UG/t2, UG/t2 zu UG/t3, KG/t1 zu KG/t2, KG/t2 zu KG/t3. Die Ergebnisse gegenüberstellend konnten wir verdeutlichen, dass UG und KG zum Messzeitpunkt t1 hinsichtlich ihres Erkrankungsgrades noch weitgehend identisch sind, dass sich dann aber zu den Messzeitpunkten t2 und t3 deutliche Unterschiede zwischen den beiden Gruppen abzeichnen, die wir mit begründeter Annahme auf die Therapieeinwirkung zurückführen. Was aber fehlt, das ist ein direktes statistisches Maß für die Unterschiede in der Entwicklung von t1 bis t3.

Darin drückt sich ein wesentlicher Nachteil des t-Tests aus: der t-Test kann immer nur zwei Datengruppen miteinander vergleichen, kann also nur Paarvergleiche rechnen, ist aber zu einer integrierten Verarbeitung komplexerer Datenbestände mit mehreren Datengruppen nicht in der Lage. Das heißt für uns, dass wir die Ergebnisse der einzelnen t-Tests zu t1, t2 und t3 zwar deskriptiv vergleichen und bewerten können, dass wir dadurch aber nicht zu einer statistisch-mathematischen Prüfgröße kommen.

Mit der Varianzanalyse lässt sich dieser Nachteil beheben, denn sie bietet eine integrierte Fehlertermbestimmung für mehr als zwei Datengruppen, macht also über einfache Paarvergleiche hinausgehende komplexere Datenvergleiche möglich. Liegen allerdings nur zwei Datengruppen vor, leistet eine Varianzanalyse nicht mehr als ein t-Test, - eine Varianzanalyse mit zwei Datengruppen ist zum t-Test rechnerisch identisch.

In unserem Fall liegen sechs Datengruppen vor: UG zu t1, UG zu t2, UG zu t3, KG zu t1, KG zu t2 und KG zu t3. Wir entschlossen uns, den Messzeitpunkt t2 aus Vereinfachungsgründen wegzulassen und uns auf den Kernbereich der Frage zu konzentrieren, nämlich ob UG und KG von t1 bis t3 einen signifikant unterschiedlichen Verlauf nehmen. Dadurch verblieben nur noch vier Datengruppen (UG zu t1, KG zu t1, UG zu t3 und KG zu t3), die es mittels Varianzanalyse, - also in einem einzigen integrierten Rechengang -, auf Signifikanz zu prüfen galt.

Wir haben uns hierzu wieder der Daten der auch bei den t-Tests herangezogenen fünf Skalen des GBB und der 13 Skalen des MMPI(K) bedient. Die Varianzanalyse erbrachte folgende Ergebnisse: In keiner Skala zeigt sich dabei ein „hoch signifikanter“ Unterschied ( $F: 0,000$ ) zwischen der Mittelwertentwicklung von UG und KG. Allerdings zeigen fünf klinische Skalen einen „sehr signifikanten“ Unterschied ( $F: 0,001$

– 0,010) zwischen der Mittelwertentwicklung von UG und KG zugunsten der UG an. Hierbei handelt es sich um die beiden GBB-Skalen Magenbeschwerden und Beschwerdedruck sowie die drei MMPI(K)-Skalen Depression, Psychasthenie und Schizoidie. „Signifikante“ Unterschiede (F: 0,011 – 0,050) in der Mittelwertentwicklung zugunsten der UG zeigen sich in vier Skalen, und zwar in der GBB-Skala Erschöpfungsneigung und den drei MMPI(K)-Skalen Lie, Hypochondrie und Psychopathie. Schließlich finden sich in einer weiteren MMPI(K)-Skala noch „tendenziell signifikante“ Unterschiede (F: 0,051 – 0,100) in der Mittelwertentwicklung, nämlich in der Skala Hysterie.

In allen anderen acht Skalen fanden sich keine signifikanten Unterschiede zwischen der Mittelwertentwicklung von UG und KG, nämlich in den GBB-Skalen Gliederschmerzen und Herzschmerzen sowie den MMPI(K)-Skalen Validity, Correction, Maskulinität/Femininität, Paranoia, Hypomanie sowie Soziale Introversion. Drei dieser „nicht signifikant“ anzeigenden Skalen sind für unsere Untersuchung - und auch klinisch - irrelevant, nämlich die MMPI(K) Skalen Validity und Correction, die lediglich MMPI(K)-test-interne Prüfgrößen darstellen, sowie die Skala Maskulinität/Femininität, deren Therapieresistenz uns nicht überraschte, sondern unter Plausibilitäts Gesichtspunkten von uns sogar eher erwartet worden war, - denn warum sollte sich die geschlechtsspezifische Identität eines Patienten durch Psychotherapie verändern, wenn diese doch in der Regel nur sehr selten ein Psychotherapieanlaß ist. Bringt man diese drei irrelevanten Skalen in Abzug, dann verbleiben nur noch fünf klinische Skalen, die in der Varianzanalyse „nicht signifikant“ angesprochen haben. Dem stehen zehn klinische Skalen gegenüber, in die mehr oder weniger deutlich die Wirkung der Psychotherapievariablen in dem Sinne zeigen, dass in der UG eine signifikant höhere Symptomminderung als in der KG nachgewiesen werden konnte.

Besonders erwähnenswert erscheint uns der Umstand, dass im MMPI(K) die drei Skalen der „neurotischen Trias“ (Hypochondrie, Depression, Hysterie) sowie die Psychasthenie-Skala zu den Skalen zählen, die signifikant angesprochen haben. Dies erscheint uns deshalb wichtig, weil diese Skalen gemäß der Mitteilung der Testautoren (vgl. Kap. 5.12.3) die zentralen Neuroseindikatoren des MMPI(K) darstellen und weil psychoanalytische Psychotherapie ja in erster Linie auf neurotische und weniger auf bekanntermaßen therapeutisch eher schwer angehbare psychosenahen Zustandsbilder oder psychopathische Persönlichkeitsstörungen abzielt. Bei den „nicht signifikant“ ansprechenden Skalen des MMPI(K) handelt es sich – unter Ausblendung der drei genannten irrelevanten Skalen - um die MMPI(K)-Skalen Paranoia, Hypomanie und Soziale Introversion. Die beiden ersten Skalen sind durchaus als eher „neurose-fern“ einzustufen, während sich die „Soziale Introversion“ zwar auch in neurotischen, - sehr gehäuft aber auch in psychosenahen Zustandsbildern findet. Von daher ist es nicht unbedingt überraschend, dass diese Skalen psychotherapeutisch/varianzanalytisch „nicht signifikant ansprechen“.

Betrachtet man die Entwicklung der Mittelwertunterschiede in den fünf relevanten aber „nicht signifikant“ ansprechenden Skalen von t1 zu t3, dann fallen trotz der fehlenden Signifikanzen auch hier die in der UG durchweg tendenziell höheren Mittelwertabsenkungen im Vergleich zu KG auf. Auffallend ist, dass sich jedoch auch bei der KG in allen Skalen Mittelwertabsenkungen finden, nur fallen diese durchweg schwächer aus als in der UG. Es zeigt sich also unterhalb des Niveaus statistischer Signifikanz auf der



deskriptiven Ebene eine stärkere Symptomverbesserung in der UG als in der KG. Für die Verbesserungen in der KG lassen sich Plausibilitätsgründe vorbringen. So könnte es sich hier entweder um Wirkungen der „Hoffnung auf Therapie“ oder um Spontanheilungen handeln. Die stärkere Absenkung der Mittelwerte in der UG lässt sich dann ebenfalls mit dem Plausibilitätsargument auf die überlegene Wirkung der Psychotherapie im Vergleich zur Spontanheilung zurückführen.

Alle diese Ergebnisse sind streng genommen jedoch nur dann richtig, wenn sichergestellt ist, dass UG und KG zum Messzeitpunkt t1 in ihrer „Krankheitsausprägung“ in etwa strukturgleich sind und zu diesem Zeitpunkt nicht bereits so stark voneinander abweichen, dass die Unterschiede zu t3 lediglich die bereits zu Beginn der Therapie bestehenden Unterschiede zeitlich versetzt erneut abbildet. Um dies zu prüfen stehen uns zwei Verfahren zur Verfügung: zum ersten zeigt uns ein Blick auf den Bartlett-Test (dieser prüft die Unterschiede bzgl. der Varianzen in zwei Stichproben zum definierten Testzeitpunkt im Sinne eines Globalvergleichs), dass es zum Messzeitpunkt noch keine signifikanten Unterschiede zwischen den Stichprobenvarianzen von UG und KG gibt (vgl. Tabellen 79 a und 79 b im Anhang), sich UG und KG also zu Beginn der Behandlung noch in keiner klinischen Skala krankheitswertig signifikant voneinander unterscheiden (der niedrigste P-Wert liegt bei 0,064). Zweitens könnten wir mit den Daten zum Messzeitpunkt t1 eine gesonderte Varianzanalyse rechnen; hierauf kann jedoch verzichtet werden, weil die Varianzanalyse zum Messzeitpunkt t1 lediglich einen Paarvergleich rechnen würde und damit identisch mit dem t-Test zum Messzeitpunkt t1 ist, der uns aber bereits vorliegt (vgl. oben Kap. 10.3.2.2, Tabelle 9) und eine weitgehende Strukturgleichheit des Störungsniveaus von UG und KG zum Messzeitpunkt t1 gezeigt hat. Bartlett-Test und t-Test bestätigen damit aussagegleich die Störungsgleichheit von UG und KG zum Messzeitpunkt t1. Wir können somit begründet annehmen, dass die mit Hilfe der Varianzanalyse aufgezeigten signifikanten Unterschiede der Entwicklungen zwischen UG und KG nicht auf einen von vorne herein vorhandenen Unterschied beider Gruppen, sondern auf die gleichsam experimentell eingeführte bzw. zurückgehaltene Psychotherapievariable zurückzuführen ist.

Es liegen nun sowohl die Ergebnisse der t-Tests wie auch das Ergebnis der Varianzanalyse vor und es liegt nahe, die Ergebnisse beider Verfahren miteinander zu vergleichen. Man kann die Ergebnisse beider Verfahren allerdings nur bedingt einander gegenüberstellen. Dies hat erstens damit zu tun, dass wie oben bereits erwähnt, die t-Tests immer nur Paarvergleiche rechnen kann, unsere Varianzanalyse hingegen vier Datenpools in einem einzigen Rechengang miteinander in ein statistisches Verhältnis setzt. Zweitens rechnet der t-Test mit höheren „N“ als die Varianzanalyse, weil er nämlich einfach auf alle „N“ zugreift, die zu einem Messzeitpunkt vorhanden sind, die Varianzanalyse hingegen alle „N“ unberücksichtigt lässt, die nicht gleichzeitig als zu beiden der gepaarten Messzeitpunkte antwortend identifiziert werden oder bei denen in der Paarzuordnung zweier Datensätze zu einer Versuchsperson Fehlantworten zu einem der beiden Messzeitpunkte ausgemacht werden können. Ein Ergebnisvergleich der beiden Prüfverfahren muss deshalb „hinken“, ist nur deskriptiv als Einschätzung und mit einer starken Unschärfe versehen möglich, erscheint aber dennoch interessant, um zu sehen, ob die durch die t-Tests vorgegebene Richtung durch die Varianzanalyse eher bestätigt oder eher in Abrede gestellt wird.

Da sich die gerechnete Varianzanalyse auf den Zeitraum t1 bis t3 bezieht, ist es nötig, ihr die im gleichen Zeitraum mit dem t-Test festgestellten Ergebnisse, also die Querschnittvergleiche der Mittelwerte von UG und KG zu den Messzeitpunkten t1, t2 und t3, gegenüberzustellen.

Die folgende Schemaübersicht liefert eine direkte Gesamtübersicht der Ergebnisse von Varianzanalyse und t-Tests:

Skala	t-Tests UG/KG			Varianzanalyse UG/KG	
	t1	t2	t3	t1:t3	
GBB-Erschöpfungsneigung	n.s.	n.s.	s.	s.	
GBB-Magenbeschwerden	n.s.	n.s.	s.	s.s.	
GBB-Gliederschmerzen	n.s.	s.	s.s.	n.s.	
GBB-Herzschmerzen	n.s.	n.s.	s.	n.s.	
GBB-Beschwerdedruck	n.s.	t.s.	s.s.	s.s.	
MMPI(K)-Hypochondriasis	n.s.	n.s.	s.	s.	Neuroseindikator
MMPI(K)-Depression	n.s.	n.s.	n.s.	s.s.	Neuroseindikator
MMPI(K)-Hysteria	n.s.	n.s.	n.s.	t.s.	Neuroseindikator
MMPI(K)-Psychasthenia	t.s.	n.s.	n.s.	s.s.	Neuroseindikator
MMPI(K)-Lie	n.s.	n.s.	n.s.	s.	
MMPI(K)-Psychopathic Deviate	s.	n.s.	n.s.	s.	
MMPI(K)-Paranoia	n.s.	n.s.	n.s.	n.s.	
MMPI(K)-Schizoidie	n.s.	n.s.	t.s.	s.s.	
MMPI(K)-Hypomanie	n.s.	n.s.	n.s.	n.s.	
MMPI(K)-Soziale Introversion	n.s.	n.s.	n.s.	n.s.	
MMPI(K)-Validity	n.s.	n.s.	n.s.	n.s.	irrelevant
MMPI(K)-Correction	n.s.	n.s.	n.s.	n.s.	irrelevant
MMPI(K)-Maskulin/Feminin	n.s.	n.s.	n.s.	n.s.	irrelevant

s.s. = sehr signifikant; s. = signifikant; t.s. = tendenziell signifikant; n.s. = nicht signifikant;

Insgesamt handelt es sich um 18 Skalen. Drei Skalen (vgl. oben) können als irrelevant außer acht gelassen werden, so dass wir von 15 aussagerelevanten Skalen ausgehen.

Zunächst die Skalen, in denen t-Tests und Varianzanalyse zu gleichgerichteten Aussagen kommen: In fünf von den insgesamt 15 relevanten Skalen weisen beide Testverfahren gemeinsam, - wenn auch mit teilweise stark voneinander abweichenden Signifikanzniveaus -, auf die Signifikanz eines unterschiedlichen Verlaufs von UG und KG hin, nämlich in den drei GBB-Skalen Erschöpfungsneigung, Magenbeschwerden und Beschwerdedruck sowie in den zwei MMPI(K)-Skalen Hypochondrie und Schizoidie.

In drei weiteren Skalen kommen beide Testverfahren ebenfalls zu übereinstimmenden Ergebnissen, allerdings in diesem Falle dazu, dass sich UG und KG „nicht signifikant“ unterschiedlich entwickelt haben. Es sind dies die drei MMPI(K)-Skalen Paranoia, Hypomanie und Soziale Introversion, - drei Skalen, die - wie oben bereits diskutiert - nicht zu den zentralen vier Neuroseindikatoren (Hypochondrie, Depression,

Hysterie und Psychasthenie) des MMPI(K) gehören und deshalb in der Gesamtschau eine mindere Relevanz haben.

Es bleiben sieben Skalen zu diskutieren, in denen eines der beiden Prüfverfahren eine Signifikanz nachweist, das andere Prüfverfahren jedoch keine Signifikanz feststellt. In zwei Skalen kommt nur der t-Test zu einem signifikanten Ergebnis, in fünf Skalen kommt lediglich die Varianzanalyse zu einem Signifikanznachweis. Im einzelnen handelt es sich um folgende Skalen:

- GBB-Gliederschmerzen: hier stellt der t-Test zum Messzeitpunkt t1 zunächst Strukturgleichheit von UG und KG fest, rechnet dann einen signifikanten Unterschied zwischen den Mittelwerten beider Gruppen zum Messzeitpunkt t2 und noch einmal einen sehr signifikanten Unterschied der Mittelwerte der beiden Gruppen zum Messzeitpunkt t3. Überraschenderweise kommt die Varianzanalyse im Vergleich der Messzeitpunkte t1/t3 zu dem Ergebnis „nicht signifikant“, obwohl der Bartlett-Test zum Messzeitpunkt t1 die Strukturgleichheit, zum Messzeitpunkt t3 hingegen den sehr signifikanten Unterschied der Mittelwerte der beiden Gruppen anzeigt. Die Erklärung dieser scheinbaren Ungereimtheit ist darin zu sehen, dass die t-Tests nur Paarvergleiche zu festen Messzeitpunkten durchführen können, dabei aber nicht die Entwicklung der Mittelwertunterschiede zu den verschiedenen Messzeitpunkten mit berücksichtigen. So kommen die t-Tests zwar zu dem Ergebnis, dass sich die Mittelwerte von UG und KG zu t2 und zu t3 absolut genommen zwar signifikant unterscheiden, es bleibt aber unberücksichtigt, dass die KG durch Spontanheilung eine zur UG tendenzgleiche Entwicklung genommen hat und so weit „nachgezogen“ hat (Abnahme des Mittelwertes in der KG von 9,658 auf 8,474), dass die recht ansehnlichen Mittelwertunterschiede zu den verschiedenen Messzeitpunkten - zu dem richtungssynchronen Verlauf beider Gruppen (Abnahme des Mittelwertes in der UG von 9,010 auf 6,375) ins Verhältnis gesetzt - lange nicht mehr so viel wiegen, als wenn die KG keine solche Entwicklung genommen hätte.

- GBB-Herzschmerzen: t-Test weist eine „einfache Signifikanz“ nach, die Varianzanalyse findet wiederum „keine Signifikanz“. Hier ist die gleiche Erklärung wie bei GBB-Gliederschmerzen zu geben: Angesichts der Mittelwertabsenkung in der KG von 6,816 (t1) auf 5,447 (t3) verliert der signifikante Mittelwertunterschied von UG (4,563) und KG (5,447) zum Messzeitpunkt t3 an Bedeutung, - bildhaft gesprochen könnte man sagen, dass der signifikante Mittelwertunterschied zum Messzeitpunkt t3 (t-Test) durch die Hinzunahme des Vergleichs der Mittelwertabsenkungen in UG und KG von t1 zu t3 wieder „aufgefressen“ wird.

- MMPI(K)-Hysterie: in dieser Skala findet der t-Test keine Signifikanz, die Varianzanalyse hingegen eine „tendenzielle Signifikanz“. Dieser Unterschied in der Aussage ist nur sehr gering und daher zu vernachlässigen.

- MMPI(K)-Depression: der t-Test findet keine Signifikanz, die Varianzanalyse findet hingegen „sehr signifikante“ Unterschiede. Dies ist ein sehr auffälliger Ergebnisunterschied. Die Erklärung liegt darin, dass die Mittelwertunterschiede von UG und KG zum Messzeitpunkt t1 wie auch zum Messzeitpunkt t3 für sich genommen (Querschnittvergleich) nicht sehr gravierend (Unterschied zu t1: 1,177; Unterschied zu t3: 1,138) sind, der Unterschied der Mittelwertentwicklung im Verlauf einer jeden Gruppe (Längsschnittvergleich) von t1 zu t3 jedoch erhebliche Differenzen ausweist. In der UG sinkt der Mittelwert von t1 zu t3 um 3,315 Punkte, in der KG dagegen nur um 1,000 Punkte. Die beiden Gruppen verhalten sich somit massiv unterschiedlich, doch wird dieser Unterschied bei der Querschnittbetrachtung

der beiden Gruppen messtechnisch nicht deutlich, - erst der Längsschnittvergleich der beiden Gruppen macht auf die Unterschiede aufmerksam, - ein Hinweis auf die überlegene Aussagekraft der Varianzanalyse im Vergleich zum t-Test. Mit dem t-Test alleine hätten wir in diesem Fall eine hochsignifikant unterschiedliche Entwicklung von UG und KG nicht erkannt.

- MMPI(K)-Lie und MMPI(K)-Psychopathie: der t-Test findet in beiden Skalen keine Signifikanz, die Varianzanalyse hingegen eine „einfache Signifikanz“. In der Lie-Skala liegen die Mittelwerte von UG und KG bezogen auf die beiden Messzeitpunkte recht eng beieinander, die divergierende Entwicklung eines leichten Mittelwertanstiegs in der UG und eines leichten Mittelwertabfalls in der KG führt jedoch im Verlauf gesehen zu einem Auseinanderdriften beider Gruppen, dem die Varianzanalyse eine Signifikanz zurechnet. In der Psychopathie-Skala rechnet der t-Test ein signifikantes Auseinanderfallen der Mittelwerte zu t1, dann gleichen sich die Mittelwerte von UG und KG zu t3 einander an, so dass sich die Signifikanz zu verlieren scheint, der Verlauf dieser Angleichung verweist jedoch wieder auf eine signifikant unterschiedliche Entwicklung von UG und KG, ohne die eine Angleichung der ursprünglich (zu t1) hinsichtlich der Mittelwerte recht unterschiedlichen Gruppen nämlich nicht möglich gewesen wäre.

- Ähnlich verhält es sich in der Skala MMPI(K)-Psychasthenie. Auch hier nähern sich die Mittelwerte von UG und KG zum Messzeitpunkt t3 bis auf 0,383 Punkte einander an. Entsprechend kommt der t-Test zu t3 zu dem Ergebnis, dass dieser Mittelwertunterschied nicht signifikant ist. Sieht man aber, dass der Mittelwert der UG zum Messzeitpunkt t1 noch bei 15,000 und bei der KG bei 12,735 lag, dann hat in der UG die bedeutend größere Mittelwertabsenkung stattgefunden. Diesen Verlauf mit berücksichtigend kommt die Varianzanalyse zu dem vom t-Test völlig abweichenden Ergebnis, dass hier ein sehr signifikant unterschiedlicher Verlauf der beiden Gruppen UG und KG stattgefunden hat.

In diesem Detailvergleich werden die unterschiedlichen Ansätze von t-Tests und Varianzanalyse deutlich. Ihre Aussagen scheinen sich zunächst zu widersprechen. Dieser Eindruck entsteht jedoch nur deshalb, weil beide Prüfverfahren unterschiedlich ausgerichtet sind: der t-Test prüft eben immer nur zwei Datengruppen, also ein Datenpaar, wohingegen die Varianzanalyse das statistische Verhältnis von mehr als zwei Datengruppen miteinander berechnen kann. Beide Verfahren sind deshalb auch nicht direkt miteinander vergleichbar. Erst die deskriptive Detailanalyse jedes einzelnen Ergebnisses macht einen Vergleich der Aussagen möglich und löst scheinbare Widersprüche auf. Dabei wird im Ganzen gesehen deutlich, dass die Varianzanalyse eine ähnliche Tendenz wie der t-Test anzeigt, nämlich dass die UG der KG hinsichtlich nachweisbarer Veränderungen im Sinne von Symptomverbesserungen deutlich überlegen ist, was wir unter den kontrollierten Bedingungen der Untersuchung so interpretieren, dass diese Überlegenheit der UG auf die Wirkung der Psychotherapie-Variable zurückzuführen ist. Danach hat die „behandelte“ UG im Ganzen gesehen signifikant bessere Symptomverbesserungen nachzuweisen als die „unbehandelte“ KG.

### 10.3.5 Indirekte Veränderungsmessung: unsere Ergebnisse und die Ergebnisse von Vergleichsgruppen

In diesem Kapitel wollen wir überprüfen, wie unsere Ergebnisse in den durchgeführten Testverfahren im Vergleich zu den Ergebnissen ausgewählter Populationen, die von den Testautoren mitgeteilt werden, einzuordnen sind. Dies ist ein weiteres wichtiges, externes, Kriterium zur Beurteilung der Wirksamkeit der von uns untersuchten Behandlungen.

#### 10.3.5.1 Die Ergebnisse im MMPI(K)

Es wurde ein *Vergleich mit Gesunden und Kranken* (s. Gehring u. Blaser, 1983) vorgenommen, um auch auf diesem Weg der Frage nach der Effektivität der individualpsychologisch-psychoanalytischen Einzeltherapie nachzugehen.

Die Rohwerte sind nachzulesen in Tabelle 3 für die Untersuchungsgruppe und in Tabelle 4 für die Kontrollgruppe.

Um unsere Werte, die als Rohwerte vorlagen, in Beziehung zu Vergleichsgruppen bei Gehring und Blaser setzen zu können, war es notwendig sie in T-Werte umzurechnen (s. Tabelle 23 im Anhang).

Zur Zusammensetzung der Patientenstichprobe von Gehring und Blaser finden sich im Handbuch folgende Angaben (S. 40):

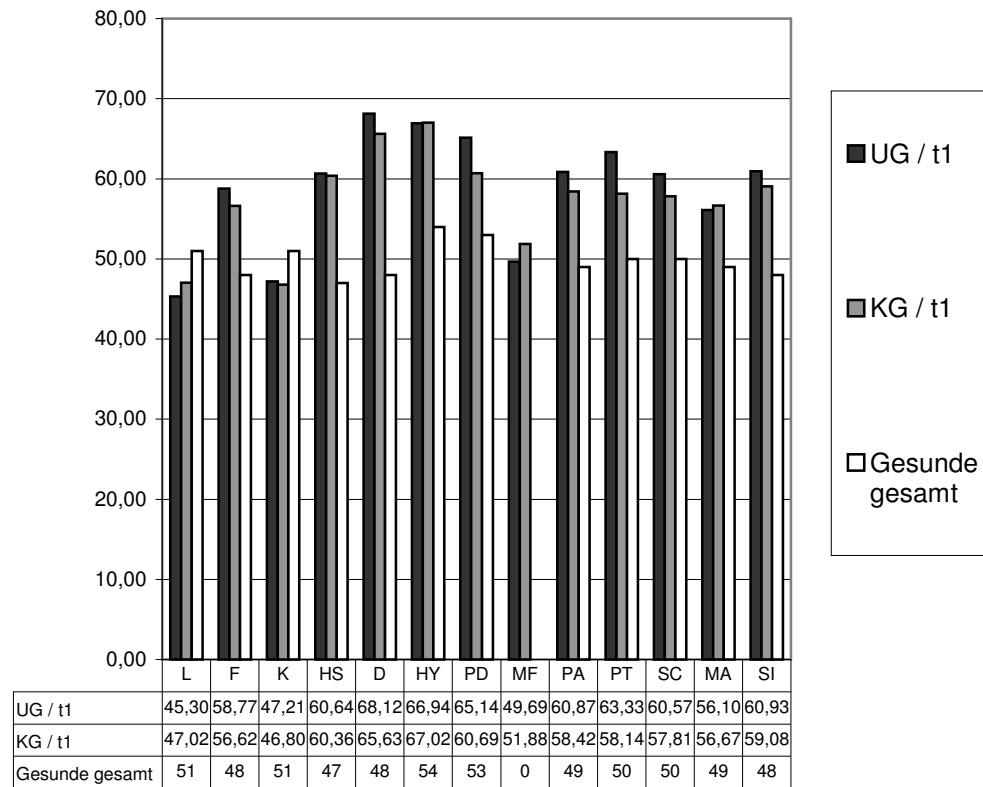
Aus dem Benutzerpool des Medical Computer Service des Verlages Hans Huber, Bern, wurden 1404 auswertbare Protokolle der MMPI-Kurzform ausgewählt. Davon stammen 42 aus Deutschland und Österreich und 1362 aus der deutschsprachigen Schweiz, 1258 aus der psychiatrischen Universitätspoliklinik Bern. Die 146 nicht aus dieser Klinik stammenden Protokolle stammen von 25 verschiedenen Psychotherapeuten und fünf psychiatrischen Kliniken.

**Tabelle 24: MMPI(K), T-Werte für die 13 Skalen: Vergleich Untersuchungsgruppe zu t1 und Kontrollgruppe zu t1 mit Gesunden, Neurotikern sowie Patienten mit Persönlichkeitsstörungen (Gehring u. Blaser, 1983)**

MMPI	UG / t1			KG / t1			Gesunde	Neurotiker		Persönlichkeitsstörungen	
Skala	Gesamt	Männer	Frauen	Gesamt	Männer	Frauen	Gesamt	Männer	Frauen	Männer	Frauen
	M	M	M	M	M	M	M	M	M	M	M
MMPI-Lie	45,30	44,67	45,57	47,02	49,00	46,25	51	59	59	58	59
MMPI-Validity	58,77	56,20	59,80	56,62	59,38	55,50	48	70	68	70	70
MMPI-Correction	47,21	46,95	47,32	46,80	46,28	47,03	51	52	52	52	51
MMPI-Hypochondriasis	60,64	60,77	60,59	60,36	67,93	57,42	47	70	69	69	66
MMPI-Depression	68,12	66,78	68,66	65,63	72,38	63,19	48	76	71	74	72
MMPI-Hysteria	66,94	66,36	67,17	67,02	73,69	64,54	54	69	68	76	66
MMPI-Psychopathie	65,14	66,68	64,54	60,69	64,75	59,06	53	70	69	70	72
MMPI-maskulin-feminin	49,69	62,00	44,79	51,88	62,77	47,94	M 52 W 48	59	50	63	52
MMPI-Paranoia	60,87	58,15	61,99	58,42	60,71	57,47	49	68	68	70	70
MMPI-Psychasthenia	63,33	69,30	60,95	58,14	69,14	53,74	50	72	67	68	67
MMPI-Schizoidie	60,57	60,72	60,50	57,81	67,54	54,09	50	76	70	77	76
MMPI-Hypomanie	56,10	55,47	56,94	56,67	59,07	55,27	49	60	60	63	61
MMPI-Soziale Introversion	60,93	63,78	59,76	59,08	66,36	56,17	48	63	62	60	62

Tabelle 24 zeigt unsere T-Werte zusammen mit den T-Werten von Vergleichsgruppen bei Gehring u. Blaser (1983). Im Vergleich zu der Gesamtgruppe „Gesunde“ liegen die Mittelwerte der Gesamtgruppe unserer UG- und KG-Patienten deutlich höher, wobei der Unterschied besonders groß bei der Skala Depression ausfällt (Abb.16). Dies kann als Krankheitsbeleg unserer Patienten interpretiert werden.

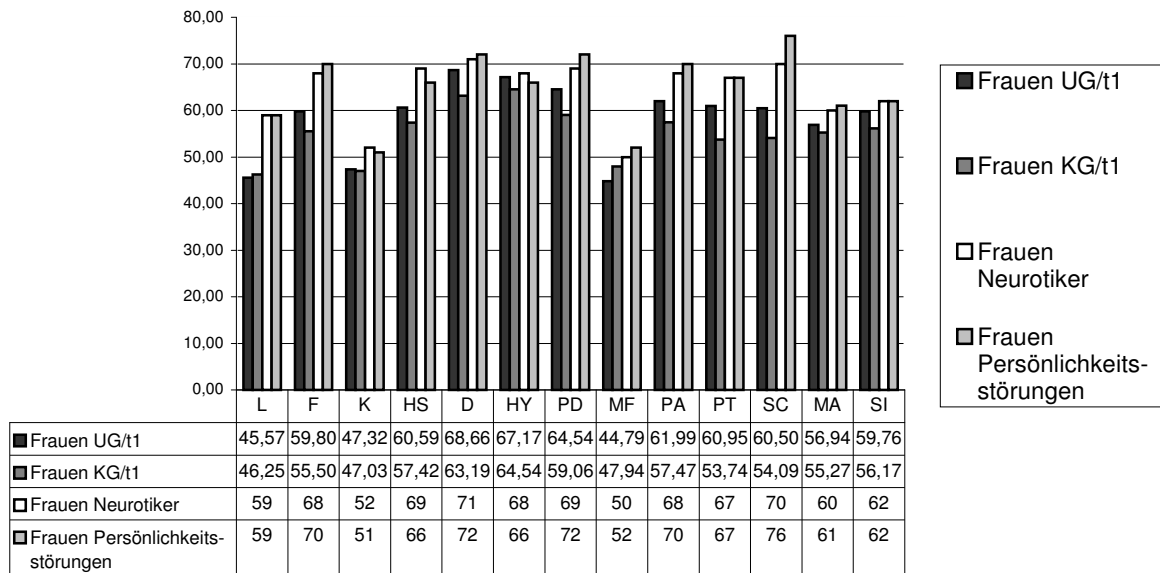
**Abb. 16: MMPI(K)-T-Werte (Gesamt): Vergleich UG und KG zu t1 mit Gesunden (Gehring u. Blaser, 1983)**



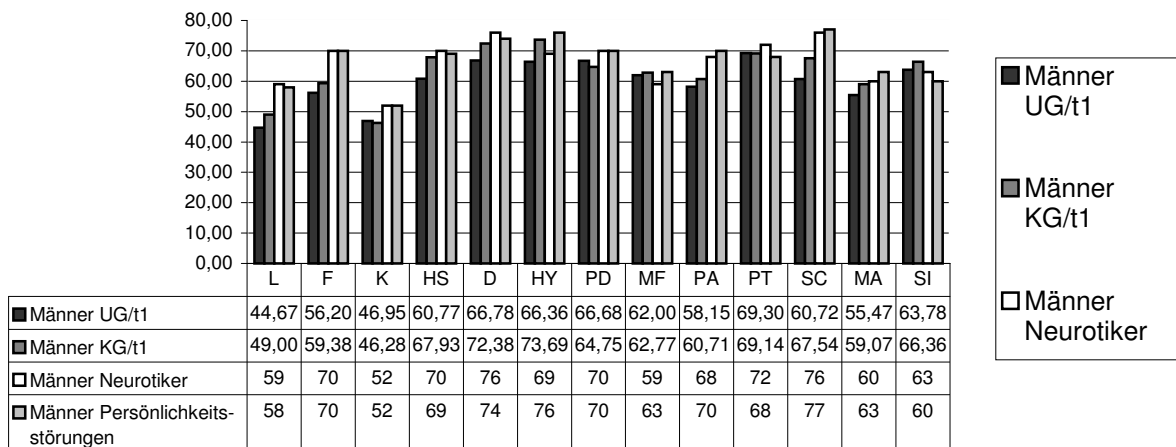
Der T-Wert für die Skala Depression in der Untersuchungsgruppe (T=68.12) liegt um rund 20 Punkte über dem der Gruppe der Gesunden bei Gehring und Blaser (T=48). Dies bedeutet, dass unsere Patienten nach diesem Kriterium depressiver sind. Beide Werte liegen jedoch noch im Durchschnittsbereich, der von T=30 bis T=70 geht, wobei unserer noch eben knapp unter der Grenze zum Überdurchschnittlichen liegt.

Im Vergleich zu der Gruppe der Neurosen (ICD 300, 8. Rev., Männer N=239, Alter: 31.7; Frauen N=295, Alter: 30.8) und der Gruppe der Persönlichkeitsstörungen (ICD 300, 8. Rev., Männer N=118, Alter: 34.0; Frauen N=220, Alter 31.6) (s. Gehring u.Blaser,1983, S.41-43), die jeweils für Männer und Frauen (als Profile) jedoch nicht für die Gesamtgruppe vorliegen, zeigen sich in unserer Untersuchungsgruppe zu meist niedrigere Werte für Männer und Frauen, was für einen geringeren Störungsgrad unserer Männer und Frauen im Vergleich zu den in der Handanweisung befindlichen Störungsgruppen sprechen würde (Abb. 17 u.18).

**Abb. 17: MMPI(K), T-Werte (Frauen): Vergleich UG und KG zu t1 mit Neurotikern und Persönlichkeitsstörungen (Gehring u. Blaser, 1983)**



**Abb. 18: MMPI(K), T-Werte (Männer): Vergleich UG und KG zu t1 mit Neurotikern und Persönlichkeitsstörungen (Gehring u. Blaser, 1983)**



Zur Erklärung der Unterschiede zwischen den Mittelwerten unserer Patienten und denen der Vergleichsgruppen bei Gehring und Blaser kann vielleicht der Schweregrad der Störungen der zum größten Teil aus der psychiatrischen Klinik stammenden Patienten herangezogen werden.

Nach der Betrachtung des Störungsgrades unserer Patienten im Vergleich zum Störungsgrad der Vergleichsgruppen soll nun die Frage geklärt werden, ob die durch die Behandlung erreichte Gesundung den Grad an Gesundheit einer gesunden Vergleichsgruppe erreicht.

**Tabelle 25: MMPI(K), T-Werte für die 13 Skalen, Vergleich der Untersuchungsgruppe zum Zeitpunkt t5 mit Gesunden (Gehring und Blaser, 1983)**

<b>MMPI-Skala</b>	<b>UG/t5 Gesamt M</b>	<b>Gesunde* Gesamt M</b>
<b>MMPI-Lie</b>	46,66	51
<b>MMPI-Validity</b>	54,05	48
<b>MMPI-Correction</b>	51,43	51
<b>MMPI-Hypochondriasis</b>	54,00	47
<b>MMPI-Depression</b>	57,08	48
<b>MMPI-Hysteria</b>	58,48	54
<b>MMPI-Psychopathie</b>	59,14	53
<b>MMPI-maskulin-feminin</b>	50,26	52/48
<b>MMPI-Paranoia</b>	55,00	49
<b>MMPI-Psychasthenia</b>	55,18	50
<b>MMPI-Schizoidie</b>	54,90	50
<b>MMPI-Hypomanie</b>	54,37	49
<b>MMPI-Soziale Introversion</b>	57,78	48

\* T-Werte aus Profil, S. 46 abgelesen.

Es ist zu erwarten, dass unsere Patienten nach dreijähriger Behandlung das Niveau der Gesunden erreichen. Dies ist jedoch, wie sich in Tabelle 25 zeigt, außer in den Kontrollskalen Lie (keine Veränderung vgl. mit t1) und Correction (Anstieg des T-Wertes vgl. mit t1) und Skala M/F (keine Veränderung zu t1) nicht der Fall. Der geringste Abstand zwischen unseren Werten und denen bei Gehring u. Blaser (1983) liegt bei ungefähr fünf Punkten (z.B. Skalen Psychasthenie, Hysterie). Woran es liegt, dass unsere Patienten doch entscheidend unter dem Niveau der Gesunden liegen, kann nicht beantwortet werden. Über die Gruppe der Gesunden heißt es bei Gehring u. Blaser lediglich, dass sie „weder neurologisch noch psychisch krank“ (S. 46) sind. Die Gruppe besteht aus 23 Frauen (Alter 28,5 Jahre) und 22 Männern (Alter 36,6 Jahre). Weitere Informationen liegen nicht vor.

Auf einen Vergleich unserer Daten mit den in der Handanweisung (Gehring u. Blaser, 1983, S. 64) befindlichen Daten zur sog. Normalpopulation haben wir aus Plausibilitätsgründen verzichtet.



Die Mittelwerte in unserer Untersuchung liegen schon zu t1 erheblich unter denen der Normalpopulation z.B. MMPI(K)-D in UG/t1: M=14.02 gegenüber M=22.2 (Männer)/23.8 (Frauen) in der Normalpopulation oder MMPI(K)-PT UG/t1: M=12,75 gegenüber M=25,5 (Männer)/29.8 (Frauen). Unsere Patienten wären demzufolge schon vor der Behandlung erheblich gesünder als die „Normalen“ bei Gehring u. Blaser. Dieser Unterschied ist nicht nachvollziehbar und wirft Fragen zu Herkunft der Daten und zu den Merkmalen der Normalpopulation auf, die leider auch in einem persönlichen Briefwechsel mit Prof. Blaser nicht geklärt werden konnten.

### Exkurs: Verfälschungstendenzen in den Antworten auf die Fragen des MMPI(K)

Hier wenden wir uns der Frage zu, ob sich *Verfälschungstendenzen* bei der Beantwortung des MMPI(K) zeigen und, wenn ja, ob sie eher zur Verschlimmerung oder zur Verharmlosung der Befindlichkeit führen. Zu diesem Zweck ziehen wir die Mittelwerte der drei Kontrollskalen des MMPI(K) Lie (L), Validity (F) und Correction (K) heran. Da der L-Wert zu t1 bei keinem Patienten über der kritischen Grenze (L=70) lag, konnte er für die weitere Berechnung unberücksichtigt bleiben. Zunächst überprüften wir Verfälschungstendenzen für die Untersuchungsgruppe, danach auch für die Kontrollgruppe. Dazu teilten wir Untersuchungsgruppe und Kontrollgruppe zu t1 jeweils in zwei Gruppen an der Grenze  $F > 70$  und  $K > 60$ . Der Mittelwert von 70 bzw. 60 stellt nach Auskunft der Testautoren jeweils die kritische Grenze dar; wird sie überschritten, so können Verfälschungstendenzen angenommen werden. Für Untersuchungsgruppe und für Kontrollgruppe haben wir nun jeweils einen Vergleich der Mittelwerte der um MMPI(K)- $F > 70$  und MMPI(K)- $K > 60$  bereinigten Gruppe mit der nicht bereinigten Gruppe vorgenommen. Dabei interessiert uns, ob es Unterschiede in den anderen klinischen Skalen des MMPI(K) zwischen den beiden Gruppen gibt.

**Tabelle 26: MMPI(K)-Skalen 1-13 (Rohwerte) zu t1, Vergleich der Mittelwerte in der Untersuchungsgruppe der um  $F > 70 + K > 60$  bereinigten Gruppe mit der nicht bereinigten Gruppe**

	bereinigt um $F > 70 + K > 60$			nicht bereinigt		
	M	S	N	M	S	N
MMPI-Lie	3,27	1,75	108	3,35	1,16	142
MMPI-Validity	3,81	2,24	110	5,01	3,87	139
MMPI-Correction	8,72	2,90	110	8,51	3,36	140
MMPI-Hypochondriasis	9,00	4,80	107	9,43	4,42	141
MMPI-Depression	13,60	4,36	109	13,00	3,39	144
MMPI-Hysteria	14,89	5,26	106	15,58	3,94	137
MMPI-Psychopathie	11,87	4,11	104	12,74	4,42	136
MMPI-maskulin-feminin	14,86	3,47	103	15,01	3,39	133
MMPI-Paranoia	7,20	3,02	104	8,18	3,94	134
MMPI-Psychasthenia	13,96	4,71	108	14,77	5,11	140
MMPI-Schizoidie	8,27	3,87	104	9,42	4,87	133
MMPI-Hypomanie	9,94	2,98	103	10,40	3,16	134
MMPI-Soziale Introversion	13,43	4,56	109	13,84	4,60	141

Tabelle 26 zeigt: Ein  $F > 70$  bzw. ein  $K > 60$  (jeweils kritische Grenze für Verfälschungstendenzen) (Gehring und Blaser, 1983) haben zum Zeitpunkt t1 bis zu 35 Patienten der Untersuchungsgruppe (ungefähr 1/4 der Grundgesamtheit), je nach Skala. Ein hoher K-Wert zeigt eine Abwehrhaltung gegenüber seelischen Schwächen, die bis zur absichtlichen Entstellung der Antworten, mit der Absicht, einen „normalen“ Eindruck zu erwecken, gehen kann. Zur Interpretation hoher F-Werte bietet sich der Grad an Gestörtheit unserer Patienten an.

Es ist nicht, wie auch im Handbuch erwähnt, von Oberflächlichkeit bei der Beantwortung oder Auffassungsproblemen auszugehen (Hathaway u. McKinley, 1991). Generell fällt auf, dass die Gruppe mit den bereinigten Werten in fast allen MMPI(K)-Skalen niedrigere Mittelwerte aufweist als die Gruppe mit den nicht-bereinigten Werten. Dies gilt allerdings nicht für die Skala Depression mit leicht erhöhtem Mittelwert. Zur Interpretation der niedrigeren Mittelwerte in der „bereinigten“ Gruppe lässt sich sagen, dass die Patienten offensichtlich dazu neigen, sich zum Zeitpunkt vor der Behandlung etwas kranker zu schildern als sie wirklich sind. Relativ stark reduzierte Mittelwerte zeigen sich logischerweise bei der Skala F, in der Tendenz aber auch bei den meisten anderen Skalen.

Auch Böhme u. Teusch (1999) finden in ihrer Untersuchung, dass Patienten, die Hilfe suchen und sich auf eine konfliktzentrierte Psychotherapie einlassen, in der Mehrzahl wenig abwehrend sind. Nur bei 25% der Patienten lässt sich nach den Befunden dieser Autoren von erhöhter Abwehr sprechen. Patienten mit der Zusatzdiagnose „Persönlichkeitsstörung“ tendierten ausgeprägter zur Verschlimmerung.

**Tabelle 27: MMPI(K)-Skalen 1-13 (Rohwerte) zu t1, Vergleich der Mittelwerte in der Kontrollgruppe der um  $F > 70 + K > 60$  bereinigten Gruppe mit der nicht bereinigten Gruppe**

	bereinigt um $F > 70 + K > 60$			nicht bereinigt		
	M	s	N	M	s	N
MMPI-Lie	4,03	1,61	33	3,74	1,80	50
MMPI-Validity	3,42	2,24	33	4,38	3,61	45
MMPI-Correction	9,00	3,01	33	8,29	3,55	45
MMPI-Hypochondriasis	9,09	4,45	33	9,52	4,53	50
MMPI-Depression	12,24	5,44	33	13,00	5,40	49
MMPI-Hysteria	14,79	5,98	33	15,46	5,73	48
MMPI-Psychopathie	9,97	4,73	31	10,79	5,18	42
MMPI-maskulin-feminin	14,09	3,08	33	14,27	3,28	49
MMPI-Paranoia	6,48	3,08	33	7,29	4,01	48
MMPI-Psychasthenia	12,16	5,72	32	13,04	6,08	49
MMPI-Schizoidie	7,09	4,18	32	8,38	5,41	47
MMPI-Hypomanie	10,47	2,95	32	10,89	3,83	46
MMPI-Soziale Introversion	11,85	5,85	33	12,86	5,73	49

In Tabelle 27 zeigt sich, wie in der Untersuchungsgruppe, für die Kontrollgruppe in allen Skalen (außer den Kontroll-Skalen) eine Abnahme der Mittelwerte, wenn die Patienten mit erhöhtem F bzw. K aus der Berechnung herausgenommen werden. Auch die Patienten der Kontrollgruppe tendieren dazu, sich als gestörter anzusehen als sie wirklich sind. Diese Gruppe macht ungefähr ein Drittel der Gesamtgruppe (Untersuchungsgruppe ein Viertel) aus. Ob der hohe Verfälschungsanteil in der Kontrollgruppe evtl. auch auf den Versuch der Patienten zurückgeht, sich über die starke Betonung ihrer Gestörtheit einen Therapieplatz vom Analytiker zu erkämpfen?

Zum Vergleich unserer Ergebnisse im MMPI(K) mit denen von Gehring u. Blaser (1983) lässt sich zusammenfassend feststellen, dass der Krankheitsnachweis für unsere Patienten zum Zeitpunkt der probatorischen Sitzungen als erbracht angesehen werden kann, auch wenn wir eine Tendenz der Patienten, sich kranker zu schildern als man ist, berücksichtigen. Die dreijährige Behandlung unserer Patienten konnte allerdings nicht die Werte der Gesunden-Gruppe bei Gehring u. Blaser erreichen, wobei der Vergleich mit der Normalpopulation ganz unplausible Ergebnisse erbringt, so dass Zweifel aufkommen, ob die Validierung unserer Ergebnisse zur Wirksamkeit der dreijährigen Behandlung anhand der „Gesunden“-Daten von Gehring u. Blaser überhaupt sinnvoll ist.

Nach dem MMPI(K) vergleichen wir im nächsten Kapitel unsere Ergebnisse im Gießener Beschwerdebogen (GBB) mit den von Brähler und Scheer mitgeteilten Mittelwerten der Eichstichprobe und der Patientengruppe (PPA).

#### **10.3.5.2 Die Ergebnisse im Gießener Beschwerdebogen (GBB)**

Wir gehen in folgenden Schritten vor:

- 1) Vergleich der Untersuchungsgruppe zu allen fünf Testzeitpunkten mit der Eichstichprobe und der Patientenstichprobe (PPA).
- 2) Vergleich von Untersuchungs- und Kontrollgruppe, jeweils zu t1 mit der Patientengruppe (PPA) sowie der Untersuchungsgruppe zu t5 mit der Eichstichprobe, jeweils auf Itemebene.
- 3) Die zehn stärksten Beschwerden in der Untersuchungsgruppe im Vergleich von t1 mit t5.
- 4) Die Differenz der zehn stärksten Beschwerden in der Untersuchungsgruppe zu t1 und zu t5 im Vergleich zu der Differenz zwischen GBB Patientenstichprobe und Eichstichprobe.
- 5) Seelische und körperliche Bedingtheit der Beschwerden, in der Untersuchungsgruppe zu t1 und zu t5, im Vergleich zu Kammerer (1998).

Die Skalenrohwerte unserer Untersuchung sind nachzulesen in Tabelle 3 für die Untersuchungsgruppe und in Tabelle 4 für die Kontrollgruppe.

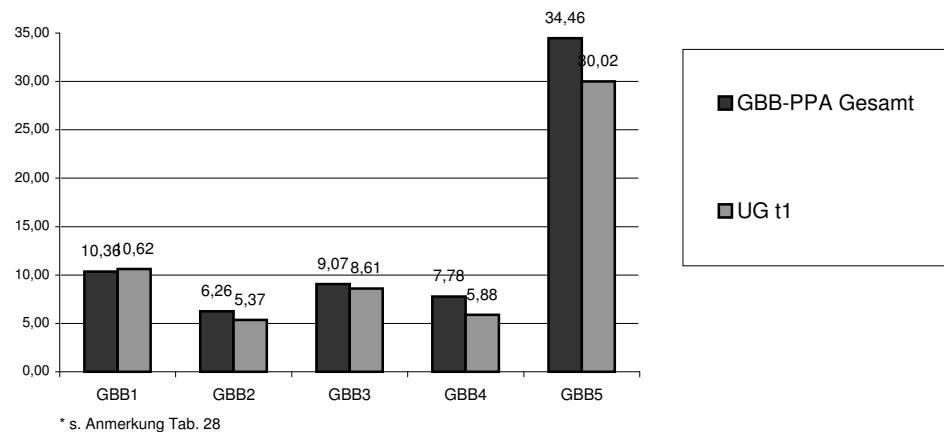
**Tabelle 28: GBB-Skalen 1 bis 5, Mittelwerte und Standardabweichungen (s)  
für GBB-Eichstichprobe, GBB-PPA und UG t1 bis t5**

	Erschöpfungs- grad		Magen- beschwerden		Glieder- schmerzen		Herz- beschwerden		Beschwerde- druck	
	M	s	M	s	M	s	M	s	M	s
GBB-Eichstichprobe										
Männer, gesamt	3,52	3,70	2,67	3,32	4,11	4,37	2,46	3,41	12,76	12,39
Frauen, gesamt	5,55	4,62	2,70	3,22	6,51	4,94	3,41	3,68	18,18	13,46
Gesamt *	4,53	-	2,68	-	5,31	-	2,93	-	15,47	-
UG t5	6,71	5,06	3,05	3,28	6,00	4,59	3,26	3,42	19,17	13,14
UG t4	6,52	5,01	2,80	3,15	5,70	4,61	3,59	3,62	18,37	12,58
UG t3	7,35	5,26	3,54	3,39	5,96	4,14	3,83	3,79	20,57	12,40
UG t2	8,77	5,74	3,93	3,55	6,89	4,58	4,58	3,65	24,09	13,25
UG t1	10,62	5,86	5,37	4,44	8,61	5,40	5,88	3,99	30,02	14,87
GBB-PPA Männer, gesamt*	9,04	-	6,11	-	7,81	-	7,19	-	31,45	-
GBB-PPA Frauen, gesamt*	11,18	-	6,20	-	10,26	-	8,38	-	37,47	-
GBB-PPA Gesamt*	10,36	-	6,26	-	9,07	-	7,78	-	34,46	-

\* von uns errechnetes arithmetisches Mittel aus Mittelwerten von Untergruppen (vgl. Brähler, 1995 S. 95, 96)

1) Der Vergleich unserer Untersuchungsgruppe zum Zeitpunkt t1 mit der Patientengruppe (gesamt) von Brähler, siehe Tabelle 28, zeigt für Skala 1 (Erschöpfungsgrad) eine recht große Übereinstimmung, tendenziell liegt der Mittelwert unserer Untersuchungsgruppe über dem der Patientenstichprobe. Dieser Befund für sich genommen spricht für einen starken Leidensdruck, in dem sich unsere Patienten vor der psychotherapeutischen Behandlung befinden. In allen anderen Skalen des GBB sind die Mittelwerte unserer Untersuchungsgruppe leicht unter den Mittelwerten der Patientengruppe (PPA) (GBB 2, Magenbeschwerden und GBB 3, Gliederschmerzen) oder deutlich darunter (GBB 4, Herzbeschwerden, GBB 5, Beschwerdedruck) (Abb.19).

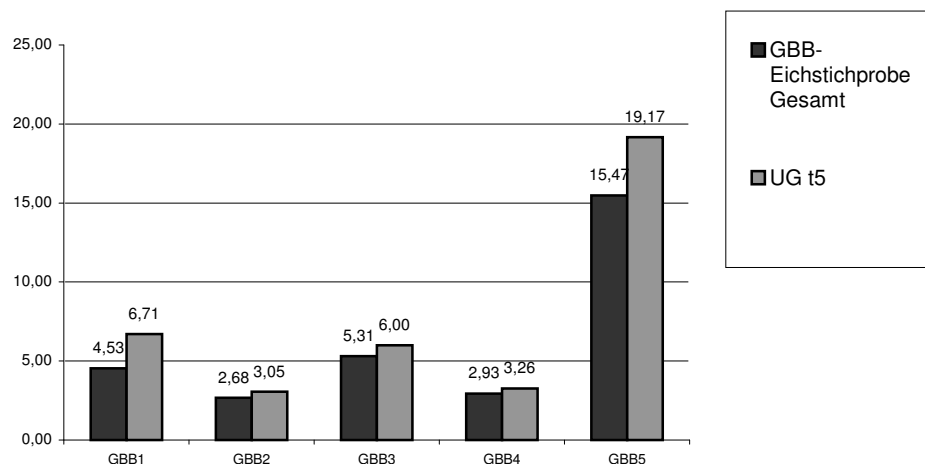
**Abb. 19: Mittelwerte für GBB-PPA Gesamt\* und UG t1**



Folgt man diesen Zahlen, so dürften unsere Patienten vor der eigentlichen psychotherapeutischen Behandlung im Grad ihrer somatoformen Störungen insgesamt etwas weniger belastet sein als die Patientengruppe von Brähler, was sich besonders in der Höhe des Beschwerdedrucks zeigt. Möglicherweise hängen diese Unterschiede auch mit der anderen demographischen Zusammensetzung unserer Patientengruppe zusammen, in der jüngere Altersgruppen und höherer Bildungsgrad überwiegen. In diesen Gruppen ist das Klageniveau (s. Brähler, 1983) im allgemeinen geringer als bei Älteren und niedrigerem Bildungsstand.

Inwieweit gesunden unsere Patienten nach dreijähriger Behandlung, so dass sie sich den „gesunden“ Patienten der Eichstichprobe annähern? Abb. 20 zeigt, dass im Grad der Erschöpfung und des Beschwerdedrucks der Abstand relativ groß ist. Magenbeschwerden, Gliederschmerzen und Herzbeschwerden haben dagegen so weit abgenommen, dass sie sich von den Gesunden der Eichstichprobe nur noch unwesentlich unterscheiden.

**Abb. 20: Mittelwerte für GBB-Eichstichprobe Gesamt \* und UG t5**



\* s. Anmerkung Tab. 28

Möglicherweise kann dieser Befund damit erklärt werden, dass die Behandlung der nach drei Jahren erfassten Patienten zum Zeitpunkt t5 noch nicht beendet war, somit die Neuroseindikatoren Erschöpfung und Beschwerdedruck noch auf einen gewissen Krankheitswert verweisen, auf Symptomebene gleichwohl schon „gesunde“ Werte erzielt werden.

Übrigens sind nach den Befunden von Kern und Küchenhoff (1998) stark körperbezogene Beschwerden zu Beginn der Behandlung ein guter Prädiktor für ein erfolgreiches Behandlungsergebnis, was sich mit unseren Zahlen bestätigen lässt.

2) Nach den Vergleichen auf Skalenebene folgt nun der Vergleich der Patientenstichprobe (PPA) (Brähler, 1983) mit Untersuchungsgruppe und Kontrollgruppe zum Zeitpunkt t1 auf der Ebene der Einzelbeschwerden (Itemebene).

**Tabelle 29: Vergleich des GBB Mittelwertprofils in der Patientenstichprobe (PPA) gesamt mit unserer Untersuchungsgruppe und Kontrollgruppe, jeweils zum Zeitpunkt t1**

Item	Kurztext	PPA N=4076		UG/ t1 N= 145		KG/ t1 N=51	
		M	s	M	s	M	s
1	Schwächegefühl	1,44	1,32	1,52	1,21	1,41	1,06
2	Herzklopfen	1,60	1,35	1,26	1,02	1,43	1,10
3	Völlegefühl	1,31	1,40	1,36	1,27	1,35	1,23
4	Neigung zum Weinen	1,60	1,42	1,91	1,26	1,59	1,20
5	Juckreiz	0,67	1,08	0,84	1,14	0,75	1,09
6	Ohnmachtsanfälle	0,23	0,68	0,13	0,43	0,09	0,30
7	Schlafbedürfnis	1,63	1,50	1,90	1,35	1,67	1,35
8	Gesch. Untererregbarkeit	0,99	1,36	1,00	1,33	1,14	1,30
9	Gliederschmerzen	0,85	1,25	0,99	1,17	1,27	1,17
10	Schwindelgefühl	1,27	1,28	1,26	1,12	1,22	1,10
11	Rückenschmerzen	1,30	1,38	1,82	1,30	1,94	1,35
12	Schwitzen	1,54	1,43	1,12	1,15	1,45	1,22
13	Nackenschmerzen	1,10	1,36	1,80	1,41	2,08	1,21
14	Gehstörungen	0,39	0,93	0,26	0,70	0,41	0,94
15	Erbrechen	0,43	0,92	0,29	0,68	0,25	0,63
16	Sehstörungen	0,71	1,06	0,59	0,95	0,73	0,96
17	Anfälle	0,29	0,86	0,14	0,56	0,27	0,76
18	Übelkeit	0,88	1,16	0,84	1,09	0,76	0,97
19	Gewichtszunahme	0,59	1,01	0,96	1,07	0,90	0,88
20	Kloßgefühl im Hals	0,97	1,30	1,10	1,24	1,27	1,13
21	Drang zum Wasserlassen	0,70	1,11	0,95	1,26	0,90	1,02
22	Hautveränderungen	0,52	1,01	0,94	1,20	1,12	1,18
23	Aufstoßen	0,92	1,20	0,75	1,02	0,71	1,06
24	Kälteüberempfindlichkeit	1,29	1,44	1,35	1,41	1,06	1,07
25	Sodbrennen	0,77	1,12	0,80	1,13	0,73	0,98
26	Verkrampfung im Arm	0,68	1,09	0,66	0,90	0,90	1,14
27	Kopfschmerzen	1,52	1,36	1,56	1,28	1,53	1,17
28	Wärmeüberempfindlichkeit	0,90	1,26	0,69	1,05	0,94	1,26
29	Erschöpfbarkeit	1,90	1,43	1,93	1,32	1,88	1,39
30	Schlafstörungen	1,66	1,49	1,80	1,24	1,68	1,24
31	Geschl. Übererregbarkeit	0,51	0,99	0,58	0,97	0,41	0,78
32	Müdigkeit	1,91	1,36	2,28	1,17	2,16	1,24
33	Gleichgewichtsstörungen	0,75	1,11	0,73	1,01	0,82	1,14
34	Schluckbeschwerden	0,33	0,78	0,37	0,84	0,41	0,75
35	Hustenreiz	0,59	0,99	0,57	0,90	0,73	0,96
36	Benommenheit	1,14	1,26	1,09	1,19	0,80	0,93
37	Taubheitsgefühl	1,25	1,31	1,08	1,15	1,00	1,13
38	Verstopfung	0,97	1,32	0,61	1,07	0,61	0,94
39	Appetitlosigkeit	0,87	1,19	0,59	0,99	0,76	1,09
40	Hitzewallungen	1,13	1,31	0,76	0,98	1,02	1,21
41	Müdigkeit in den Beinen	1,18	1,32	1,08	1,25	0,84	1,01
42	Mattigkeit	1,66	1,32	1,92	1,21	1,63	1,26
43	Durchfälle	0,49	0,90	0,75	1,00	0,82	1,18
44	Lähmungen	0,12	0,55	0,07	0,31	0,20	0,60
45	Stiche in der Brust	1,07	1,24	0,90	1,04	0,96	1,02
46	Zittern	1,34	1,30	0,86	0,95	0,75	0,87
47	Halsschmerzen	0,47	0,86	0,59	0,84	0,86	0,98
48	Leichtes Erröten	1,19	1,27	1,19	1,17	1,25	1,16
49	Kalte Füße	1,60	1,49	1,70	1,52	1,61	1,31
50	Heißhunger	0,72	1,12	1,30	1,29	1,06	1,10
51	Magenschmerzen	1,09	1,26	1,32	1,25	1,24	1,12
52	Atemnot	0,71	1,18	0,66	1,00	0,71	1,10
53	Unterleibsschmerzen	0,63	1,07	0,87	1,06	0,82	1,09
54	Gewichtsabnahme	0,61	1,13	0,39	0,84	0,39	0,70
55	Druckgefühl im Kopf	1,23	1,38	1,47	1,38	1,31	1,30
56	Herzbeschwerden	0,94	1,23	0,69	1,01	0,67	0,95
57	Sprachstörung	0,50	0,93	0,48	0,83	0,65	0,84

Tabelle 29 macht deutlich, dass bei fast 24 Beschwerden (von insgesamt 57) die Patienten unserer Untersuchungsgruppe zu t1 höhere Mittelwerte haben, bei elf Beschwerden sind die Unterschiede marginal bzw. nicht vorhanden, bei 22 Beschwerden sind die Mittelwerte in der Patientenstichprobe (PPA) größer als in unserer Untersuchungsgruppe zum Zeitpunkt vor der Behandlung. Für die Kontrollgruppe ergibt sich PPA < KG: 23 Beschwerden, PPA > KG: 18 Beschwerden, PPA =/~ 16 Beschwerden, d.h. die Relation ist der in der Untersuchungsgruppe sehr ähnlich.

Im nächsten Schritt folgt nun der Vergleich der GBB - Eichstichprobe mit den Mittelwerten unserer Untersuchungsgruppe zum Zeitpunkt t5 auf der Ebene der Einzelbeschwerden (Itemebene) (Tabelle 30).

Hier geht es darum festzustellen, ob unsere Patienten nach dreijähriger Behandlung den Grad an Gesundheit, wie er in der Gesamtbevölkerung besteht, erreicht.

**Tabelle 30: Vergleich des GBB Mittelwertprofils in der Eichstichprobe und unserer UG zu t5**

Item	Kurztext	GBB- Eichstichprobe		UG/ t5	
		M	s	M	s
1	Schwächegefühl	0,64	0,96	1,01	1,01
2	Herzklopfen	0,87	1,13	0,71	0,79
3	Völlegefühl	0,72	1,05	0,92	0,94
4	Neigung zum Weinen	0,71	1,09	1,08	0,92
5	Juckreiz	0,38	0,83	0,66	0,82
6	Ohnmachtsanfälle	0,08	0,35	0,03	0,16
7	Schlafbedürfnis	1,02	1,25	1,19	1,17
8	Gesch. Untererregbarkeit	0,35	0,73	0,71	0,73
9	Gliederschmerzen	0,83	1,17	0,81	1,08
10	Schwindelgefühl	0,69	1,02	0,70	0,95
11	Rückenschmerzen	1,26	1,28	1,34	1,18
12	Schwitzen	1,01	1,24	0,81	1,13
13	Nackenschmerzen	0,85	1,19	1,44	1,20
14	Gehstörungen	0,29	0,81	0,15	0,46
15	Erbrechen	0,14	0,45	0,08	0,28
16	Sehstörungen	0,56	1,00	0,32	0,64
17	Anfälle	0,05	0,34	0,07	0,30
18	Übelkeit	0,27	0,63	0,41	0,66
19	Gewichtszunahme	0,76	1,05	0,59	0,80
20	Kloßgefühl im Hals	0,28	0,70	0,60	0,97
21	Drang zum Wasserlassen	0,32	0,74	0,68	0,94
22	Hautveränderungen	0,24	0,68	0,63	0,94
23	Aufstoßen	0,50	0,87	0,44	0,75
24	Kälteüberempfindlichkeit	0,76	1,16	0,78	1,07
25	Sodbrennen	0,52	0,91	0,49	0,90
26	Verkrampfung im Arm	0,34	0,75	0,34	0,65
27	Kopfschmerzen	1,12	1,19	0,96	0,95
28	Wärmeüberempfindlichkeit	0,59	1,02	0,48	0,94
29	Erschöpfbarkeit	0,70	1,02	1,21	1,14
30	Schlafstörungen	0,70	1,06	0,96	1,10
31	Geschl. Übererregbarkeit	0,32	0,67	0,29	0,68
32	Müdigkeit	1,22	1,12	1,62	1,08
33	Gleichgewichtsstörungen	0,38	0,78	0,25	0,64
34	Schluckbeschwerden	0,13	0,44	0,32	0,76
35	Hustenreiz	0,37	0,76	0,37	0,72
36	Benommenheit	0,32	0,68	0,58	0,91
37	Taubheitsgefühl	0,67	1,04	0,55	0,83
38	Verstopfung	0,55	1,02	0,30	0,72
39	Appetitlosigkeit	0,21	0,58	0,14	0,38
40	Hitzewallungen	0,56	1,02	0,52	0,97
41	Müdigkeit in den Beinen	0,75	1,10	0,71	0,99
42	Mattigkeit	0,71	0,97	1,11	1,05
43	Durchfälle	0,17	0,51	0,40	0,81
44	Lähmungen	0,06	0,37	0,09	0,34
45	Stiche in der Brust	0,41	0,80	0,53	0,80
46	Zittern	0,28	0,66	0,45	0,82
47	Halsschmerzen	0,35	0,68	0,29	0,61
48	Leichtes Erröten	0,50	0,87	0,66	0,93
49	Kalte Füße	0,85	1,24	1,04	1,28
50	Heißhunger	0,54	0,92	0,73	0,93
51	Magenschmerzen	0,54	0,93	0,66	0,96
52	Atemnot	0,28	0,73	0,30	0,74
53	Unterleibsschmerzen	0,36	0,77	0,64	0,95
54	Gewichtsabnahme	0,20	0,58	0,21	0,64
55	Druckgefühl im Kopf	0,59	0,97	0,74	0,93
56	Herzbeschwerden	0,45	0,88	0,41	0,78
57	Sprachstörung	0,08	0,37	0,25	0,55



Tabelle 30 zeigt den Vergleich der GBB-Eichstichprobe mit den Mittelwerten unserer Untersuchungsgruppe zum Zeitpunkt t5 auf Itemebene. In 28 Fällen sind die Werte unserer Untersuchungsgruppe höher als die der Eichstichprobe, in 14 Fällen sind die Mittelwerte der Eichstichprobe größer als die Mittelwerte unserer Untersuchungsgruppe, in 15 Fällen sind die Unterschiede marginal bzw. nicht vorhanden.

In  $\frac{3}{4}$  aller Einzelbeschwerden konnte also in unserer Untersuchungsgruppe nach dreijähriger Behandlungszeit das Gesundheitsniveau der Normalbevölkerung mindestens erreicht, wenn nicht überschritten werden. Dies ist ein überzeugender Nachweis für die Wirksamkeit der von uns untersuchten Behandlungen in Bezug auf die somatoformen Beschwerden wie sie im GBB gemessen werden.

3) Als nächstes überprüfen wir, wie sich die Hauptbeschwerden durch die Behandlung verändern. Diese Frage ist von besonderem Interesse, weil die Patienten unter den stärksten Beschwerden auch am stärksten leiden, und deshalb auch ein dringender Linderungsbedarf besteht. Wir betrachten die zehn stärksten Beschwerden unserer Untersuchungsgruppe und ihre Entwicklung vom Anfang bis zum Ende der Behandlung und vergleichen Ausgangs- und Endniveau mit der Patienten- und Eichstichprobe von Brähler, 1983 (Tabelle 32).

**Tabelle 31: Rangreihe der zehn häufigsten Beschwerden im GBB in unserer UG zu t1 und t5**

t1		t5	
1	Müdigkeit (2,28)	1	Müdigkeit (1,62)
2	Erschöpfbarkeit (1,93)	2	Nackenschmerzen (1,44)
3	Mattigkeit (1,92)	3	Rückenschmerzen (1,34)
4	Neigung zum Weinen (1,91)	4	Erschöpfbarkeit (1,21)
5	Schlafbedürfnis (1,90)	5	Schlafbedürfnis (1,19)
6	Rückenschmerzen (1,82)	6	Mattigkeit (1,11)
7,5	Nackenschmerzen (1,80)	7	Weinen (1,08)
7,5	Schlafstörungen (1,80)	8	Kalte Füße (1,04)
9	Kalte Füße (1,70)	9	Schwächegefühl (1,01)
10	Kopfschmerzen (1,56)	10	Schlafstörung (0,96)

Tabelle 31 zeigt die zehn Beschwerden, unter denen unsere Patienten zum Zeitpunkt der probatorischen Sitzungen (t1) am stärksten leiden. Nach dreijähriger Behandlung (t5) verändert sich die Rangfolge der Beschwerden; bis auf Kopfschmerzen, ersetzt durch Schwächegefühl, bleiben es jedoch dieselben. Befunde aus der Mannheimer Kohortenstudie (Rudolf, 1998) zeigen eine hohe Verlaufsvarianz somatoformer Erkrankungen. Wie in unserer Studie kommt es zu einem hohen Prozentsatz zum Übergang in eine andere psychogene Erkrankung. Allerdings bleibt bei uns der Pool der Einzelstörungen unverändert, nur der Ausprägungsgrad der Einzelstörungen wechselt.

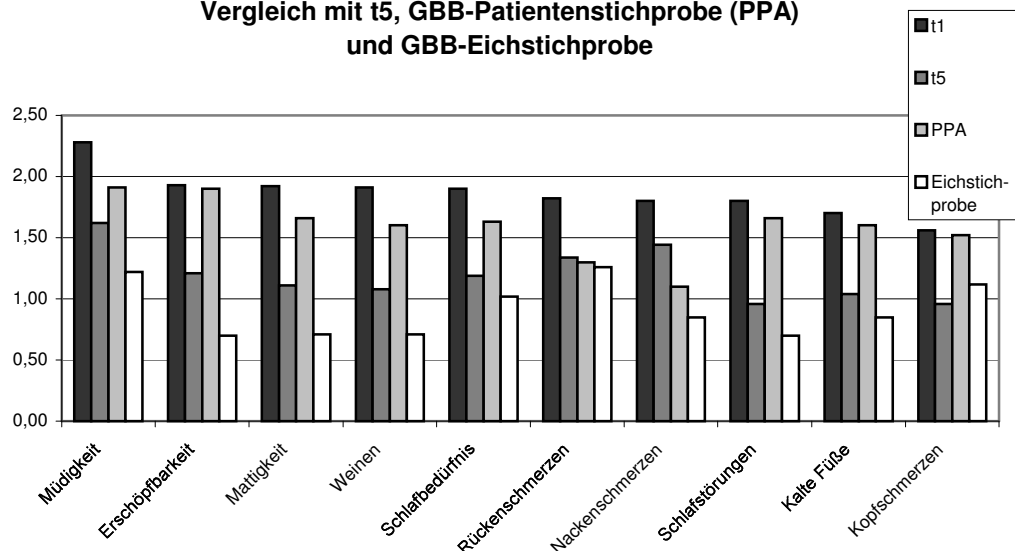
**Tabelle 32: Die zehn häufigsten Beschwerden (Mittelwerte) in unserer UG zu t1 und t5, ihre Differenzen im Vergleich zur Differenz zwischen GBB Patientenstichprobe (PPA) und Eichstichprobe (Brähler, 1983)**

		t1	t5	Differenz	PPA	Eichstich- probe	Differenz
1	Müdigkeit	2,28	1,62	0,66	1,91	1,22	0,69
2	Erschöpfbarkeit	1,93	1,21	0,72	1,90	0,70	1,20
3	Mattigkeit	1,92	1,11	0,81	1,66	0,71	0,95
4	Weinen	1,91	1,08	0,83	1,60	0,71	0,89
5	Schlafbedürfnis	1,90	1,19	0,71	1,63	1,02	0,61
6	Rückenschmerzen	1,82	1,34	0,48	1,30	1,26	0,04
7	Nackenschmerzen	1,80	1,44	0,36	1,10	0,85	0,25
8	Schlafstörungen	1,80	0,96	0,84	1,66	0,70	0,96
9	Kalte Füße	1,70	1,04	0,66	1,60	0,85	0,75
10	Kopfschmerzen	1,56	0,96	0,60	1,52	1,12	0,40

4) Die zehn häufigsten Beschwerden in unserer Untersuchungsgruppe zu Behandlungsbeginn (t1) sind: Müdigkeit, Erschöpfbarkeit, Mattigkeit, Weinen, Schlafbedürfnis, Rückenschmerzen, Nackenschmerzen, Schlafstörungen, kalte Füße, Kopfschmerzen.

Unsere Patienten haben bei diesen Beschwerden ein stärkeres Krankheitsgefühl als die Patienten bei Brähler. Insgesamt ist bei unseren Patienten das Ausgangsniveau bei allen zehn Beschwerden höher als bei der PPA. Die Werte der Eichstichprobe konnten demzufolge in unserer UG nach dreijähriger Behandlung (t5) nur bei Kopfschmerzen erreicht bzw. unterschritten werden. In den anderen neun Beschwerden erreichen unsere Patienten nicht den Grad an "Gesundheit" der Eichstichprobe. Bei Betrachtung der Differenz zeigen sich bei vier Beschwerden (5, 6, 7, 10) größere Unterschiede zwischen t5 und t1 als zwischen Eichstichprobe und Patientenstichprobe, bei vier (2, 3, 8, 9) umgekehrt und bei zwei Beschwerden (1, 4) sind die Differenzen nahezu gleich (siehe Abb. 21).

**Abb. 21: Die zehn häufigsten Beschwerden in der UG zu t1 im Vergleich mit t5, GBB-Patientenstichprobe (PPA) und GBB-Eichstichprobe**



Das heißt, gemessen an der „Wegstrecke“, die ein Patient von „Krankheit“ zu „Gesundheit“ zurücklegt, erzielen die Patienten unserer Untersuchungsgruppe bei der Hälfte der Beschwerden das Maß an Heilungsfortschritt, das auch der gedachte Patient der PPA bis zu den „gesunden“ Werten der Eichstichprobe erreicht.

Der Vergleich unserer Werte mit denen von Brähler macht möglicherweise eine Zunahme bestimmter Beschwerden im Laufe der letzten 20 Jahre deutlich. Dies trifft für die Beschwerden Müdigkeit, Mattigkeit, Weinen, Schlafbedürfnis und ganz besonders für Rückenschmerzen und Nackenschmerzen zu. Hier scheint der Beschwerde- oder Leidensdruck erheblich zugenommen zu haben, so dass bei den meisten Beschwerden die „gesunden“ Werte von vor 20 Jahren nicht erreicht werden konnten. Vielleicht ist es erlaubt, aus diesen Zahlen auf eine größere Morbidität, die in Richtung verstärkter Depressionsentwicklung weist, zu schließen.

5) Auch an der Einschätzung der körperlichen bzw. seelischen Bedingtheit der Beschwerden und ihrer Entwicklung bis zum Beobachtungsende nach drei Jahren lässt sich möglicherweise die Effektivität der Behandlung feststellen, wenn man annimmt, dass die Einschätzung seelischer Bedingtheit der Beschwerden vom Behandlungsbeginn bis zum -ende zu-, die der körperlichen Bedingtheit dagegen abnimmt.

**Tabelle 33: Vergleich der Antworten auf die Frage, ob die Beschwerden eher körperlich oder eher seelisch bedingt sind, gemessen zu den Zeitpunkten t1 bis t5 in der UG (Prozente auf der Basis der jeweils Antwortenden)**

	t1		t2		t3		t4		t5	
	seelisch	körperlich	seelisch	körperlich	seelisch	körperlich	seelisch	körperlich	seelisch	körperlich
	%	%	%	%	%	%	%	%	%	%
nicht	0,0	15,0	0,0	16,4	0,0	14,3	1,2	22,4	0,0	11,3
etwas	3,4	41,4	4,8	36,9	9,0	37,8	8,3	31,6	6,9	32,4
zum Teil	9,7	32,1	21,8	35,2	26,1	39,8	27,4	32,9	27,8	45,1
überwiegend	40,7	9,3	37,9	9,0	33,3	7,1	39,3	10,5	40,3	11,3
allein	46,2	2,1	35,5	2,5	31,5	1,0	23,8	2,6	25,0	0,0
<b>Gesamt</b>	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Tabelle 33: Bei der Frage nach der seelischen bzw. körperlichen Bedingtheit ihrer Beschwerden sind zum Zeitpunkt der Probesitzungen t1 mehr als 4/5 der Patienten davon überzeugt, dass ihre Beschwerden überwiegend oder ausschließlich seelisch bedingt sind, zu t5 (nach dreijähriger Behandlung) sind es etwas mehr als 2/3. Von der körperlichen Bedingtheit ihrer Beschwerden sind dagegen zu t1 und zu t5 etwas mehr als 10% der Patienten überzeugt. Dass die Beschwerden nicht oder nur etwas körperlich bedingt sind, meinen zu t1 über 50%, zu t5 etwas mehr als 40%. Interessant ist vielleicht auch noch die Entwicklung der Kategorie „zum Teil“: während bei der Einschätzung der seelischen Bedingtheit ihrer Beschwerden zu t1 hier knapp 10% ankreuzen, sind es zu t1 bei der körperlichen Bedingtheit knapp 1/3. Zu t5 kreuzen bei der Frage nach der seelischen Bedingtheit knapp 28% „zum Teil“ an, bei der Frage nach der körperlichen Bedingtheit sind es rund 45 %.

Es scheint sich hier eine Tendenz zu einer größeren Ausgewogenheit in der Beurteilung der seelischen/körperlichen Bedingtheit der Beschwerden im Laufe der Behandlung zu zeigen. Es ist also nicht so wie angenommen, dass die Patienten am Anfang der Behandlung überwiegend davon ausgehen, dass ihre körperlichen = Körperbezogenen Beschwerden körperlich bedingt sind, und dass dieser Anteil zum Ende der Behandlungszeit zugunsten einer seelischen Bedingtheit der Beschwerden abnimmt, sondern sie gehen schon davon aus – und das ist ja auch wahrscheinlich mit ausschlaggebend für die Wahl einer analytischer Behandlung - dass seelische Gründe für ihre Beschwerden vorliegen. Nach dreijähriger Behandlungszeit wächst dann der Anteil der körperlichen Gründe für die Beschwerden, vielleicht als Zeichen für einen gewachsenen Realismus im Laufe der Behandlung, vielleicht aber auch als Möglichkeit, den ausgebliebenen Erfolg der Behandlung der seelischen Gründe für die Beschwerden zu rationalisieren.

In Ermangelung von Vergleichsdaten zur Einschätzung der körperlichen und seelischen Bedingtheit der Beschwerden bei Brähler und Scheer haben wir im folgenden unsere Mittelwerte mit denen von Kammerer (1998) verglichen.

**Tabelle 34: Antworten auf die Frage, ob die Beschwerden eher körperlich oder eher seelisch bedingt sind; Vergleich der Ergebnisse von Kammerer (1998) mit unserer Untersuchungsgruppe**

	Beschwerden sind eher körperlich bedingt				Beschwerden sind eher seelisch bedingt			
	Anfang		Ende		Anfang		Ende	
	K %	* %	K %	** %	K %	* %	K %	** %
nicht	11,0	15,0	13,0	11,3	4,0	0,0	4,0	0,0
etwas	25,0	41,4	25,0	32,4	6,0	3,4	9,0	6,9
zum Teil	46,0	32,1	43,0	45,1	34,0	9,7	30,0	27,8
überwiegend	17,0	9,3	17,0	11,3	48,0	40,7	50,0	40,3
allein	2,0	2,1	2,0	0,0	7,0	46,2	7,0	25,0

**K** in Kammerer-Studie  
(am Anfang bzw. Ende der Behandlung)  
**\*** in unserer Studie t1  
**\*\*** in unserer Studie t5

In Tabelle 34 werden unsere Ergebnisse bzgl. der körperlichen/seelischen Bedingtheit der Beschwerden mit den Ergebnissen der Kammerer-Studie (1998) verglichen. In dieser Studie ist am Anfang der Behandlung die Überzeugung größer als in unserer Studie, dass die Beschwerden eher körperlich bedingt sind, entsprechend ist in unserer Studie am Anfang die Überzeugung von der eher seelischen Bedingtheit der Beschwerden größer. Möglicherweise spielt hier die Unterschiedlichkeit der Patienten eine Rolle, die bei Kammerer stationär behandelt werden, das heißt auch kranker sind als unsere Patienten. Während sich bei Kammerer die Überzeugung zur seelischen/körperlichen Bedingtheit der Beschwerden bis zum Ende kaum verändert, zeigt sich in unserer Untersuchung – wie gesehen - eine gewisse Tendenz zu mehr kör-

perlicher Mitbedingtheit von t1 bis t 5 und entsprechend weg von der ausschließlichen seelischen Bedingtheit der Beschwerden.

*Zusammenfassung* der Befunde aus dem GBB: Der Vergleich mit der Patientenstichprobe bei Brähler zeigt, dass unsere Patienten zum Zeitpunkt der Probesitzungen tatsächlich psychosomatisch krank sind, wenn auch nicht ganz in dem Ausmaß wie die Patienten bei Brähler. Nimmt man die Einzelbeschwerden, so zeigt sich im Vergleich zu den Ergebnissen von Brähler ein guter Therapieerfolg (bei 75% Gesundheit bis auf Normalniveau oder darüber). Bei den zehn stärksten Beschwerden in unserer Untersuchungsgruppe konnte bei einem durchweg, z. T. erheblich höheren Ausgangsniveau als bei den Patienten bei Brähler (1983) die Normalwerte der Brählerschen Eichstichprobe nur bei Kopfschmerzen erreicht werden. Auf Skalenniveau betrachtet ist der Therapieerfolg (Erreichen der „Normalen“) bei drei Skalen (Magenbeschwerden, Gliederschmerzen und Herzbeschwerden) evident, bei zweien (Erschöpfung und Beschwerdedruck) nicht. In der Einschätzung der seelischen/körperlichen Bedingtheit der Beschwerden zeigen unsere Patienten von Anfang bis zum Ende der Beobachtungszeit eine vergleichsweise höhere Veränderung zugunsten der körperlichen Mitbedingtheit ihrer Beschwerden.

Im folgenden Kapitel stellen wir unseren Mittelwerten im VEV Vergleichswerte von Zielke u. Kopf-Mehnert gegenüber.

### 10.3.5.3 Die Ergebnisse im VEV

Wie stellen sich die Veränderungen im VEV (s. Kap. 10.3.1) im Vergleich zu den Gruppen bei Zielke u. Kopf-Mehnert (1978) dar?

Zur Beantwortung dieser Frage stellen wir den VEV–Gesamtwert zu t2 und t3 für unsere Untersuchungs- und Kontrollgruppe der Therapie- und Wartegruppe von Zielke u. Kopf-Mehnert (1978) gegenüber.

**Tabelle 35: VEV - UG und KG zu t2 und t3 im Vergleich zu Zielke  
Therapie- und Wartegruppe**

	<b>Zielke*</b> <b>Therapie-</b> <b>gruppe</b>	<b>Zielke**</b> <b>Warte-</b> <b>gruppe</b>	<b>UG</b> <b>t2</b>	<b>KG</b> <b>t2</b>	<b>UG</b> <b>t3</b>	<b>KG</b> <b>t3</b>
M	222,22	171,51	185,16	171,56	195,73	173,18
S	34,53	45,72	28,31	32,46	37,06	38,83

\* Nach 9 Wochen Therapie

\*\* Nach 9 Wochen Wartezeit, Zielke (1978) S. 34, 35

Tabelle 35: Der Vergleich mit den Daten von Zielke u. Kopf-Mehnert (1978, S. 85) ergibt folgendes Bild: Der Mittelwert der Zielke-Therapiegruppe (nach neun Wochen Therapie) liegt über denen unserer Untersuchungsgruppe zu t2 (ein halbes Jahr in Behandlung) und zu t3 (ein Jahr in Behandlung), der Mittelwert der Zielke-Wartegruppe entspricht unseren Kontrollgruppenwerten zu t2 und t3. Es zeigt sich, dass die Kontrollgruppe in den VEV-Werten stagniert, die Untersuchungsgruppe dagegen Veränderungen aufweist, die sich dem Wert der Zielke-Therapiegruppe zwar nähern, aber auch noch nicht zu t5 (nach dreijähriger Behandlung) erreichen ( $M=206.42$  – s. Tabelle 5 – gegenüber  $M=222.2$ ). Es ist jedoch zu berücksichtigen, dass unser Wert ebenso wie der der Zielke-Therapiegruppe im 0,1 % Signifikanzbereich liegt.

Zur Interpretation der unterschiedlich schnellen Entwicklung der Zielke-Therapiegruppe und unserer UG muss berücksichtigt werden, dass sich die Gruppe der Klienten bei Zielke und Kopf-Mehnert aus Probanden zusammensetzt, die wegen neurotischer Störungen in der Beratungsstelle des Psychologischen Instituts der Universität Hamburg um therapeutische Einzelgespräche nachsuchten, also vermutlich weniger krank waren als unsere Patienten und auch als jüngere weniger Behandlungszeit benötigen (siehe auch unsere Ergebnisse Kap. 10.4.3).

### **10.3.6 Globalmaß der Therapieeffektivität (Effektstärke)**

Als weitere Möglichkeit, die Effektivität der von uns untersuchten Therapien in Zahlen auszudrücken, wählten wir ein Globalmaß, in dem die wichtigen Einzelmaße aggregiert sind. Die Ermittlung der Effektstärke wird auch von Langenmayr u. Kosfelder (1998) bei großen Stichproben empfohlen. Die Effektstärkenberechnung in der Psychotherapieforschung geht zurück auf die Meta-Analyse von empirischen Forschungsarbeiten (Glass, 1976, zit. bei Wittmann u. Matt, 1986, S. 20 f.), in der eine Integration und Kumulation statistischer Kennwerte versucht wird. Bei der Meta-Analyse nach Glass wird ein Maß für die Effektstärke verwendet, das aus der standardisierten Mittelwertsdifferenz zwischen Experimental- und Kontrollgruppe besteht (s. auch Kap. 2, Smith, Glass u. Miller, 1980).

Es werden heute sehr unterschiedliche Berechnungsmethoden verwendet. Eine kurze Zusammenfassung unterschiedlicher Verfahren zur Effektstärkenberechnung in Metaanalysen findet sich bei Tschuschke et al., 1997, S. 41.

Wir verwenden zur Beschreibung des Therapieeffektes ein kombiniertes Maß, in dem sowohl Beschwerdeverringerungen (GBB–Skala 5, Beschwerdedruck, die die Skalen 1–4 des GBB umfasst) als auch die Veränderungen in den wichtigen Skalen der neurotischen Trias (Hypochondrias, Depression, Hysterie) + Psychasthenie des MMPI(K) berücksichtigt werden. Den VEV haben wir nicht mit einbezogen, weil zu t1 noch keine Daten vorliegen und wir ein einheitliches Maß für alle Zeitpunkte haben wollten. Die Berechnung des entsprechenden Maßes ist angelehnt an Groß u. Riedel (1995, S. 58), die sich wiederum an Grawe (1994) orientieren. Auch Kazdin (1994, S.55) gibt dieses Maß zur Berechnung von intraindividuellen Veränderungen an. Die Veränderungen in den einzelnen Maßen werden für jeden Patienten durch die entsprechende Standardabweichung dividiert. Die ungleich stärkere Gewichtung einzelner Maße durch

skalenbedingte Abweichungen in den Differenzen wird dadurch verhindert. Die so gewonnenen standardisierten Einzelmaße werden addiert und durch die Gesamtzahl einbezogener Maße (= 5) dividiert.

### Formel für die Berechnung des Therapieeffektes

$$\text{Therapieeffekt} = \frac{\frac{GBB5(m D Z1-Z2)}{s(Z1-Z2)} + \frac{HS(m D Z1-Z2)}{s(Z1-Z2)} + \frac{D(m D Z1-Z2)}{s(Z1-Z2)} + \frac{HY(m D Z1-Z2)}{s(Z1-Z2)} + \frac{PT(m D Z1-Z2)}{s(Z1-Z2)}}{5}$$

Z1 = Zeitpunkt 1  
Z2 = Zeitpunkt 2

m D = jeweilige mittlere Differenz  
s = jeweilige Standardabweichung

Im einzelnen berechnen wir nach dieser Formel eine Reihe unterschiedlicher Maße:

- 1) Therapieeffekte in der Untersuchungsgruppe
- 2) Therapieeffekte in der Kontrollgruppe
- 3) Therapieeffekt in der Subgruppe „Therapeutenberuf“
- 4) Therapieeffekt in der Subgruppe „erfahrene vs. unerfahrene Therapeuten“

#### 1) Therapieeffekte in der Untersuchungsgruppe

**Tabelle 36: Therapieeffekt\* für verschiedene Zeitspannen  
in der Untersuchungsgruppe**

#### UG Therapieeffekt

	t1:t2	t1:t3	t1:t4	t1:t5	t2:t3	t3:t4	t3:t5
<b>N</b>	101	89	65	57	85	65	59
<b>M</b>	0,4666	0,6403	0,9088	0,8190	0,3208	0,1977	0,3045
<b>Med</b>	0,4719	0,5909	0,8925	0,7773	0,2367	0,1149	0,2918
<b>s</b>	0,8003	0,8666	0,8850	0,9549	0,7612	0,8348	0,8524

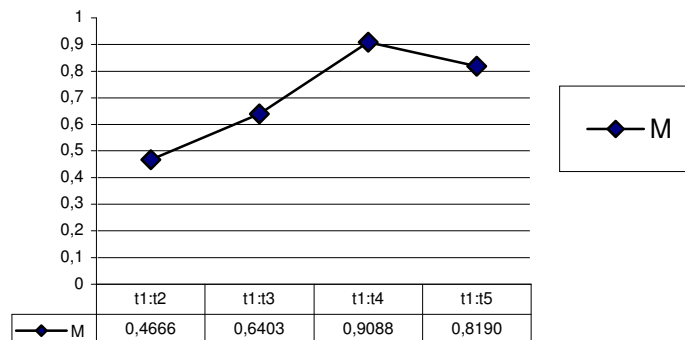
\* Der Therapieeffekt wird ausgedrückt durch den Mittelwert (M), berechnet nach der obigen Formel. Je größer M, desto größer ist der Effekt.

In Tabelle 36 sind für die Untersuchungsgruppe einige Effektstärken für den Zeitraum von t1 bis t5 zusammengestellt.

Vom Zeitpunkt der probatorischen Sitzungen (t1) aus betrachtet, erhöht sich der Therapieeffekt über den Zeitpunkt t2 (nach ½ Jahr), t3 (nach einem Jahr), bis zum Zeitpunkt t4 (nach zwei Jahren Behandlung), um dann bis zum Zeitpunkt t5 (nach drei Jahren Behandlung) leicht zurückzugehen (s. auch Abb. 22). Auf der Grundlage der Befunde von Smith, Glass u. Miller (1980), die im Mittel eine Effektstärke von .85 für

psychotherapeutische Behandlungen erwarten lassen, erreichen unsere Behandlungen eine ausreichend gute Wirksamkeit.

**Abb. 22: Therapieeffekte in der UG**



Die Beiträge der einzelnen Zeitabschnitte zum Therapieeffekt sind unterschiedlich. Der Zuwachs an Effektivität nimmt mit zunehmender Behandlungsdauer ab. Dies lässt sich durch die aufgeführten Maße t1:t2, t2:t3, t3:t4 gut zeigen. Der Effektivitätszuwachs im Zeitraum t3:t5 (ein Zweijahreszeitraum nach einem Jahr Behandlung) entspricht ungefähr dem Zuwachs für den Zeitraum t2:t3 (zweites Halbjahr nach Beginn). Die Ergebnisse bestätigen die Befunde von Howard et al. (1986): Mit fortschreitender Behandlung sind immer mehr Sitzungen nötig, um Veränderungen zu erzielen.

Aus diesen Forschungsergebnissen haben Grande et al.(1997) „den Schluss gezogen, dass die spezifischen Wirkungen von Psychoanalysen, die sich erfahrungsgemäß erst nach einer langen und intensiven Behandlung einstellen, mit konventionellen Untersuchungsinstrumenten nicht erfasst werden können“ (S. 416).

## 2) Therapieeffekte in der Kontrollgruppe

**Tabelle 37: Therapieeffekt\* für verschiedene Zeitspannen in der Kontrollgruppe**

KG Therapieeffekt					
	t1:t2	t1:t3	t1:t5 +	t3:t4 ++	t3:t5 +
<b>N</b>	31	31	18	20	14
<b>M</b>	0,1464	0,2005	0,5101	0,4967	0,4732
<b>Med</b>	0,1555	0,1997	0,5089	0,4043	0,5505
<b>Sta</b>	0,7121	0,8508	0,7395	0,6854	0,7542

\* Der Therapieeffekt wird ausgedrückt durch den Mittelwert (M), berechnet nach der Formel auf Seite 146. Je größer M, desto größer der Effekt.

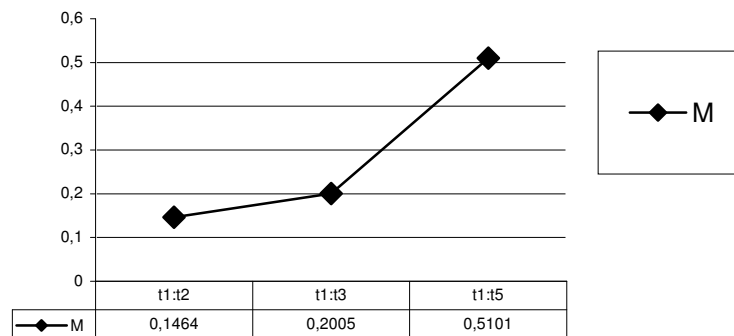
+ in der KG bedeutet t5 zwei Jahre in Therapie (nach einem Jahr Wartezeit)

++ in der KG bedeutet t4 ein Jahr in Therapie (nach einem Jahr Wartezeit)



Tabelle 37 enthält Maße für die Effektstärke von Nicht-Behandlung in der Kontrollgruppe für den Zeitraum t1 bis t3. Die Zahlen demonstrieren einen sehr geringen Effekt in der einjährigen Wartezeit und bleiben erwartungsgemäß erheblich unter den Maßen der Untersuchungsgruppe für denselben Zeitraum zurück (s. auch Abb. 23).

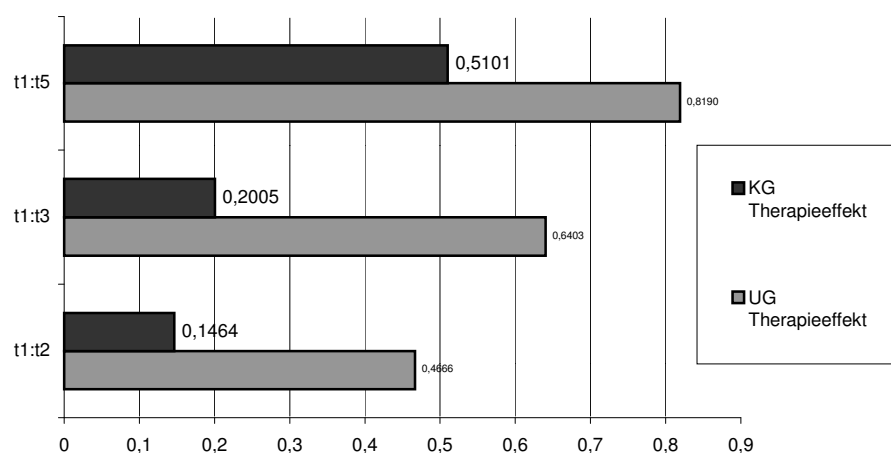
**Abb. 23: Therapieeffekte in der KG**



Interessant ist die Entwicklung des Therapieeffekts in der Kontrollgruppe, nachdem die Patienten die Behandlung nach einem Jahr Wartezeit begonnen haben. Die Patienten erreichen im ersten Jahr der Behandlung einen Effekt, der durch das nachfolgende Behandlungsjahr nicht gesteigert werden kann. Die einjährige Wartezeit hat auf das Ergebnis der nachfolgenden zweijährigen Behandlung keinen nachweisbaren positiven Einfluss.

In der eigentlichen Behandlungszeit erreichen die Kontrollgruppen-Patienten nicht annähernd den Effekt der Untersuchungsgruppe, weder im ersten noch im zweiten Behandlungsjahr - nach dem einen Jahr Wartezeit - verglichen mit dem ersten und zweiten Behandlungsjahr der Untersuchungsgruppe (Abb. 24).

**Abb. 24: Therapieeffekte in Kontrollgruppe und Untersuchungsgruppe**



Vollkommen unterschiedlich ist die Effektivität nach zweijähriger Behandlungszeit, im Vergleich von Wartegruppe (t3:t5) und Untersuchungsgruppe (t1:t4). Hat das einjährige Warten die Patienten der Kontroll-

gruppe demoralisiert, so dass die Behandlung bei ihnen weniger anschlägt? Oder befinden sich in der Kontrollgruppe doch die schwerer zu behandelnden Patienten? Auch nach jeweils einjähriger Behandlungszeit ist der Effekt in der UG (t1:t3) größer als in der KG (t3:t4). Über den gesamten Beobachtungszeitraum von drei Jahren (t1:t5) betrachtet, kann die Wartegruppe das erste Wartejahr ohne Behandlung gegenüber dem ersten Behandlungsjahr der Untersuchungsgruppe nicht kompensieren.

Die beiden folgenden Tabellen zeigen die Verteilung der Effektstärken in Quartilsabständen für Untersuchungsgruppe (UG) und Kontrollgruppe (KG).

**Tabelle 38: Therapieeffekt (t1-t5) in der UG,  
Effektstärke in Quartilsabständen**

Quartil	N	Effektstärke	obere Grenze	untere Grenze
1. Quartil	14	2,00	3,34	1,31
2. Quartil	15	1,06	1,23	0,78
3. Quartil	14	0,52	0,75	0,28
4. Quartil	14	-0,33	0,26	-1,83
1. - 3. Quartil	43	1,19	3,34	0,28

Diese Tabelle 38 zeigt die Effektstärke der gesamten dreijährigen Behandlungszeit (t1–t5) in der Untersuchungsgruppe in Quartilsabständen. Im ersten Quartil (ein Viertel der UG-Patienten) wird eine durchschnittliche Effektstärke von 2,00, im zweiten Quartil von 1,06, im dritten Quartil von 0,52 erreicht, im Durchschnitt der drei Quartile (75 % der UG-Patienten) von 1,19, im vierten Quartil zeigt die Effektstärke eine Verschlechterung der Befindlichkeit. Das heißt  $\frac{3}{4}$  der drei Jahre behandelten Patienten erzielen einen beachtlichen therapeutischen Gewinn.

Ansatzpunkt für weitere Forschung sind die in unserem Datenmaterial nicht herauszufindenden Gründe für die Verschlechterung von  $\frac{1}{4}$  der Patienten.

**Tabelle 39: Therapieeffekt (t1 - t3) in der KG,  
Effektstärke in Quartilsabständen**

Quartil	N	Effektstärke	obere Grenze	untere Grenze
1. Quartil	8	1,10	1,47	0,83
2. Quartil	7	0,48	0,69	0,30
3. Quartil	8	-0,09	0,21	-0,31
4. Quartil	8	-0,74	-0,32	-1,43

In der Tabelle 39 ist die Effektstärke des einjährigen Wartens (t1-t3) in der Kontrollgruppe in Quartilsabständen aufgeführt. Im ersten Quartil wird immerhin eine durchschnittliche Effektstärke von 1,10 des Wartens erreicht, das heißt allein durch Warten bzw. durch den Effekt der Probesitzungen kann bei einem Viertel der Kontrollgruppe eine Verbesserung in den gemessenen Skalen nach einem Jahr festgestellt werden. Diese vielleicht nicht ganz echte Remissionsquote, weil ja schon therapeutische Kontakte stattgefunden haben, zeigt sich auch in den Befunden von Franz et al. (1998), wo bei 28,4 % der zuvor somatoform erkrankten Probanden zum Zeitpunkt der Folgestudie keine ICD-Diagnose einer psychogenen Erkrankung mehr vergeben wurde.

Zur Verbesserung der Ergebnisse in der Wartezeit bieten sich folgende Erklärungen an: Schon in den wenigen Probesitzungen und auch durch die Gewissheit des baldigen Therapiebeginns bei einem Analytiker, zu dem der Patient Vertrauen gefasst hat, werden Veränderungsprozesse in Gang gesetzt. Ein einziges Beurteilungsgespräch (assessment interview) kann schon für Besserung sorgen, wenn ein heilsamer Kontakt zustande gekommen ist (s. Bergin & Lambert, 1978, S. 148).

Im zweiten Quartil zeigen sich nur noch sehr geringe Verbesserungen, die Hälfte der Kontrollgruppen-Patienten (03. und 04. Quartil) verändert sich nicht bzw. verschlechtert sich.

3) Im folgenden interessierte uns die berufspolitisch brisante Frage, ob ärztliche oder psychologische Psychoanalytiker (DGIP) einen unterschiedlich großen Therapieeffekt erzielen.

**Tabelle 40: Therapeutenberuf und Therapieeffekt (t1-t5)**

<b>Beruf</b>	<b>Therapieeffekt</b>	<b>N</b>
Arzt	0,80	11
Psychologe	0,83	45

Tabelle 40 zeigt das unspektakuläre und vielleicht beruhigende Ergebnis, dass der Therapeutenberuf Arzt oder Psychologe keine diskriminative Wirkung auf den Therapieeffekt (t1-t5) hat. Die beiden Koeffizienten sind nahezu gleich groß.

4) Als nächstes wenden wir uns der Frage zu, ob erfahrenere Therapeuten einen höheren Therapieeffekt bei ihren Patienten erreichen als unerfahrenere. Erfahrungheit wird von uns mit der Länge der Zeit nach Abschluss der analytischen Weiterbildung (Prüfungsjahr) in einem anerkannten Weiterbildungsinstitut der DGIP operationalisiert (s. Tabelle 41).

**Tabelle 41: Prüfungsjahr der an der Untersuchung beteiligten Therapeuten/innen**

		Häufigkeit	Prozent	Gültige	Kumulierte
Gültig	64	1	1,8	2,0	2,0
	77	1	1,8	2,0	4,0
	78	2	3,6	4,0	8,0
	79	1	1,8	2,0	10,0
	80	1	1,8	2,0	12,0
	83	4	7,3	8,0	20,0
	84	4	7,3	8,0	28,0
	85	2	3,6	4,0	32,0
	86	6	10,9	12,0	44,0
	87	3	5,5	6,0	50,0
	88	2	3,6	4,0	54,0
	89	4	7,3	8,0	62,0
	91	7	12,7	14,0	76,0
	92	7	12,7	14,0	90,0
	93	2	3,6	4,0	94,0
	94	2	3,6	4,0	98,0
	95	1	1,8	2,0	100,0
	<b>Gesamt</b>	50	90,9	100,0	
	<b>System</b>	5	9,1		
<b>Fehlend</b>					
<b>Gesamt</b>		55	100,0		

N	Gültig	50
M		87,06
Med.		87,50
Varianz		31,77

**Tabelle 42: t-Test der Mittelwertgleichheit des Therapieeffektes (t1:t5) von „erfahrenen“ und weniger „erfahrenen“ Therapeuten**

	Therapeuten- erfahrung	N	M	T	df	Sig. 2seit.
Therapieeffekt t1:t5	erfahren	17	0,8444	0,084	52	0,934
	weniger erfahren	37	0,8682			

Tabelle 42 zeigt, dass der Therapieeffekt der „erfahrenen“ und der „weniger erfahrenen“ Therapeuten/innen, gemessen für den gesamten Untersuchungszeitraum von drei Jahren nahezu identische Mittelwerte ergibt. Unsere Vermutung, die „alten Hasen“ könnten deutlich bessere Behandlungsergebnisse erzielen als die weniger Erfahrenen, konnte nicht bestätigt werden. Dies mag daran liegen, dass unsere über dem Median vollzogene Trennung (s. Tabelle 41) nicht scharf genug zwischen dem „Mehr oder Weniger“ an Erfahrung differenziert oder auch das Kriterium Abschlussjahr der analytischen Ausbildung nur ein unzureichendes Kriterium für die therapeutische Erfahrung darstellt.

Vielleicht würde eine differenzierte Gruppenbildung, z.B. genauere Abstufung der Erfahrung in mehrere Erfahrungsgruppen vom gerade fertigen bis hin zum Therapeuten mit bis zu 20jähriger Erfahrung andere Ergebnisse liefern.

In der Literatur werden unterschiedliche Ergebnisse berichtet. Marziali et al. (1981) konnten belegen, dass ausschließlich der Beitrag der Patienten zur therapeutischen Allianz eine Voraussagekraft bezüglich des Therapieerfolgs hatte und nicht der Beitrag des Therapeuten.

Die Ergebnisse der als *Vanderbilt-Psychotherapie-Projekt* bekannt gewordenen Untersuchung (Strupp) bestätigen ebenfalls den großen Einfluss der Patienten - Faktoren wie Ich-Stärke, Organisationsstruktur, Reife, Motivation und die Fähigkeit zum produktiven Sich-Einlassen in der Therapie.

Zu anderen Ergebnissen kommt Sandell (1999), der größere Erfolge von älteren erfahreneren im Vergleich zu weniger erfahreneren Psychoanalytikern konstatierte. „Ausgesprochen günstig schienen höheres Lebensalter sowie langjährige Berufserfahrung in der niedergelassenen Praxis nach der Zulassung als Therapeut. Es gibt einen fast einheitlichen Trend dahingehend, dass ältere Therapeuten bessere Ergebnisse erzielen, unabhängig vom Geschlecht des Therapeuten und des Patienten und unabhängig davon, ob es sich um analytische und psychodynamische Psychotherapie oder Psychoanalyse handelt" (Sandell, 1999, S. 339).

## 10.4 Varianzanalysen

Nach den Effektstärkenberechnungen zu verschiedenen Zeitpunkten für die Gesamtheit der sich in der Untersuchungsgruppe oder Kontrollgruppe befindlichen Patienten geht es nun um die Frage, ob und wie sich die Wirksamkeit der psychotherapeutischen Behandlung in verschiedenen in der Literatur immer wieder diskutierten Untergruppen unserer Stichprobe (Untersuchungsgruppe) im Behandlungsverlauf (Zwei-Punkt-Vergleich unterschiedlicher Zeitpunkte) entwickelt, und ob es Unterschiede innerhalb der Gruppen gibt.

So wollen wir klären, ob folgende globale Annahmen bestätigt werden können:

- 1) Art und Schwere der Erkrankung (Diagnose),
- 2) das Geschlecht der Patienten,
- 3) das Alter der Patienten,
- 4) die Zusammensetzung der Dyade Therapeut(in) – Patient(in),
- 5) die Sitzungsfrequenz,
- 6) Therapeutengeschlecht,
- 7) Therapeutenberuf,
- 8) Schulbildung des Patienten,
- 9) Höhe des Anfangs-Lie (MMPI[K]),
- 10) Anfänglicher Beschwerdedruck,
- 11) Anfangsdepression

beeinflusst den Therapieprozess so, dass sich unterschiedliche Entwicklungen in Zwei-Punkt-Messungen zu den unterschiedlichen fünf Zeitpunkten (von t1 bis t5) in unseren Tests (MMPI[K], GBB und VEV) statistisch nachweisen lassen. Beim MMPI(K) haben wir uns auf die vier wichtigen Skalen, die als Neuroseindikatoren gelten (Hathaway u. McKinley, 1991, S. 48) beschränkt: Neurotische Trias, i.e. Hypochondriasis (HS), Depression (D), Hysterie (HY) sowie Psychasthenie (PT).

Um die Annahmen zu überprüfen, haben wir die in der folgenden Zusammenstellung aufgeführten Varianzanalysen mit den Daten aus der Untersuchungsgruppe gerechnet.

**Tabelle 43: Zusammenstellung der gerechneten Varianzanalysen  
mit je zwei unabhängigen Variablen  
(Subgruppen und Messzeitpunkte)**

Veränderung von Zeitpunkt 1 bis Zeitpunkt 2

<b>Subgruppen</b>	<b>t1-t3</b>	<b>t3-t4</b>	<b>t3-t5</b>	<b>t4-t5</b>	<b>t1-t5</b>
Diagnosegruppen	x		x		x
Patientengeschlecht	x		x		x
Patientenalter	x		x		x
Dyade Therapeut-Patient	x		x		x
Stundenfrequenz zu t3	x				
zu t4		x			
zu t5			x	x	
Therapeutengeschlecht	x		x		x
Therapeutenberuf	x		x		x
Schulabschluss	x		x		x
Lie-Skala (t1)	x		x		x
Beschwerdedruck (t1)	x		x		x
Depression (t1)	x		x		x

#### **10.4.1 Varianzanalysen mit den Diagnosegruppen (lt. ICD-10) „neurotische Störung“, „affektive Störung“, „spezifische Persönlichkeitsstörung“ im Zeitverlauf von t1 bis t3, t3 bis t5 und t1 bis t5 und den Mittelwerten der GBB-Skalen 1 bis 5, der MMPI(K)-Skalen HS, D, HY, PT und des VEV**

1) Vergleichszeitraum t1 bis t3 (s. Tabelle 44 im Anhang):

Alle Diagnosegruppen (nach ICD-10) haben zu t1 ein ähnlich hohes Ausgangsniveau, es gibt keine signifikanten Unterschiede zwischen den Gruppen. Das heißt, dass unsere Testskalen keine Unterschiede zwischen diesen Gruppen abbilden.

In der Entwicklung von t1 bis t3 zeigt sich ebenfalls in keiner der zur Überprüfung verwendeten Skalen ein signifikanter Unterschied zwischen den drei Diagnosegruppen. Es fällt jedoch auf, dass die Patienten mit der Diagnose „neurotische Störungen“ zu t3 in MMPI(K)-Hypochondrie und in MMPI(K)-Depression niedrigere Werte haben als die Patienten mit den anderen Diagnosen. Die „Neurotiker“ scheinen tendenziell etwas rascher von der Behandlung zu profitieren als die Patienten mit den anderen Diagnosen. Auffälliger noch ist die relativ geringe Abnahme der Werte für die Persönlichkeitsstörungen in MMPI(K)-Depression, MMPI(K)-Hysterie und MMPI(K)-Psychasthenie. Die Diagnosen „neurotische Störung“ und „affektive Störung“ (depressive Neurose) erreichen bei nahezu gleichen Ausgangswerten zu t1 niedrigere Werte zu t3. Obgleich die Unterschiede nicht signifikant sind, kann doch von einer gewissen Tendenz gesprochen werden, dass die Persönlichkeitsstörungen im ersten Behandlungsjahr nicht die, jedenfalls nicht mit unseren Skalen nachweisbaren, offensichtlichen Veränderungen zeigen.

Eine ähnlich unterschiedliche Entwicklung lässt sich in den GBB-Skalen (vielleicht mit der Ausnahme von GBB-Erschöpfung) nicht feststellen.

## 2) Vergleichszeitraum t3 bis t5 (s. Tabelle 45 im Anhang):

Zum Zeitpunkt t3 unterscheiden sich zwar die Ausgangswerte der Diagnosegruppen, jedoch nicht signifikant (mit Ausnahme GBB-Herzbeschwerden).

Die Entwicklung der Diagnosegruppen von t3 bis t5 macht Unterschiede deutlich, die bei GBB-Gliederschmerzen und bei MMPI(K)-Psychasthenie signifikant sind: Bei den affektiven Störungen (depressive Störung) nimmt die Neigung zu Zwangsverhalten und Phobien (MMPI[K]-Psychasthenie) sowie Gliederschmerzen relativ deutlich ab, während sich in den beiden anderen Diagnosegruppen keine oder nur geringfügige positive Veränderungen zeigen.

In den anderen Skalen der neurotischen Trias treten Unterschiede zutage zwischen „neurotische Störungen“ einerseits und „affektive Störungen“ bzw. „spezifische Persönlichkeitsstörungen“ andererseits. Bei letzteren gehen die Werte zurück, bei ersteren stagnieren sie, dies gilt auch für den Beschwerdedruck (GBB 5), der sich bei den neurotischen Störungen im Beobachtungszeitraum von zwei Jahren im Gegensatz zu den anderen beiden Diagnosegruppen nicht mehr verändert, während diese ihren schon zu t3 geringeren Beschwerdedruck zum Ende des Beobachtungszeitraumes noch einmal vermindern. Dieselbe Entwicklung zeigt sich bei der Veränderungsmessung mit Hilfe des VEV. Zwar sind auch hier die Unterschiede zwischen den Diagnosegruppen nicht signifikant, Patienten mit affektiver Störung und insbesondere spezifischer Persönlichkeitsstörung gewinnen am Ende des dritten Behandlungsjahres jedoch ein „Mehr“ an Entspannung und Gelassenheit hinzu und gelangen in die Nähe des Niveaus der Patienten mit neurotischer Störung am Ende des ersten Behandlungsjahres.

## 3) Vergleichszeitraum t1 bis t5 (s. Tabelle 46 im Anhang):

Auf der Basis der reduzierten Grundgesamtheit von Patienten mit vollständigen Daten zu t1 und t5 zeigen sich zu t1 keine signifikanten Unterschiede zwischen den drei Diagnosegruppen (so auch Tabelle 42), obwohl in den GBB-Skalen die Mittelwertunterschiede zwischen „neurotische Störung“ und „affektive Störung“ einerseits und „spezifische Persönlichkeitsstörungen“ andererseits auffallen. In allen GBB-Skalen bleiben die Mittelwerte für die Persönlichkeitsstörungen unter denen der beiden anderen Gruppen, während in den vier MMPI(K)-Skalen (neurotische Trias +Pt) die Unterschiede weniger stark ausgeprägt sind. Dies könnte ein Hinweis darauf sein, dass psychosomatische Störungen zu Beginn der Behandlung bei den Persönlichkeitsstörungen eine vergleichsweise geringere Rolle spielen als bei den neurotischen und affektiven (i.e. Depression) Störungen. Insgesamt bleibt das (mit unseren Verfahren abgebildete) Störungsniveau der Persönlichkeitsstörungen zum Zeitpunkt t1 eher unter dem der anderen beiden Gruppen. Möglich ist, dass die schwerer Gestörten in der Zwischenzeit die Behandlung beendet hatten.

Auch die Veränderung der drei Diagnosegruppen bis zum Ende des dritten Behandlungsjahres macht gewisse Unterschiede in der Entwicklung deutlich. In der Tendenz profitieren nach diesen Zahlen die Depressiven (affektive Störung) (Diagnosestellung zu t1) am besten von der Behandlung, besonders in GBB-Beschwerdedruck und MMPI(K)-Psychasthenie, wo die Unterschiede tendenziell signifikant bzw. signifikant sind, aber auch in fast allen anderen Skalen.



Sind unsere Kollegen „Spezialisten“ für die Behandlung von Depressionen? Oder brauchen die anderen Störungsgruppen evtl. noch eine längere Zeit (die wir ja nicht mehr erfasst haben) zur Gesundung?

*Zusammenfassung der Ergebnisse:* Die drei Diagnosegruppen unterscheiden sich zu Behandlungsbeginn kaum in der Schwere der Störung, tendenziell scheinen die schwerer gestörten Patienten mit der Diagnose „spezifische Persönlichkeitsstörung“, die Behandlung nicht bis zum Ende des dreijährigen Beobachtungszeitraumes geführt zu haben. Es scheint, dass im ersten Behandlungsjahr die Gruppen „neurotische Störung“ und „affektive Störung“ (depressive Neurose) mehr von der Behandlung profitieren als die Gruppe „Spezifische Persönlichkeitsstörung“. Im folgenden zweijährigen Behandlungsabschnitt sind es dann die spezifischen Persönlichkeitsstörungen und wiederum die Patienten mit der Diagnose „affektive Störung“, die mehr von der Behandlung profitieren als Patienten mit der Diagnose „neurotische Störung“. Nimmt man den gesamten Beobachtungszeitraum von drei Jahren, so machen die Patienten mit der Diagnose „affektive Störung“, in der sich ausschließlich die mit der Ursprungsdiagnose „depressive Neurose“ versehenen Patienten befinden, die günstigste Entwicklung.

#### **10.4.2 Varianzanalysen mit Patientengeschlecht im Zeitverlauf von t1 bis t3, t3 bis t5, t1 bis t5 und den Mittelwerten der GBB-Skalen 1 bis 5, der MMPI(K)-Skalen HS, D, HY, PT und des VEV**

##### **1) Vergleichszeitraum t1 bis t3 (s. Tabelle 47 im Anhang):**

Zum Zeitpunkt t1 gibt es außer in GBB-Gliederschmerzen und tendenziell in MMPI(K)-Hysterie keine signifikanten Unterschiede zwischen Männern und Frauen.

Im Zeitverlauf von t1 bis t3 ist der Unterschied zwischen Männern und Frauen statistisch nicht signifikant. Das heißt, Frauen klagen zu Beginn der Behandlung mehr über Gliederschmerzen und entwickeln mehr konversionsneurotische Symptome als Männer. Der durch die Behandlung erfolgte Rückgang der Symptome hat für beide Geschlechter keinen auffällig verschiedenen Verlauf, tendenziell vielleicht am ehesten noch in der Depressionsentwicklung, wo die Frauen von einem etwas höheren Niveau ausgehend sich zu mehr relativer Depressionsfreiheit verändern.

##### **2) Vergleichszeitraum t3 bis t5 (s. Tabelle 48 im Anhang):**

Zu t3 sind die Unterschiede zwischen Männern und Frauen in allen betrachteten Skalen außer dem VEV nicht signifikant. Im VEV sind die Unterschiede zwischen Männern und Frauen zu t3 sehr signifikant, Frauen haben einen deutlich höheren Veränderungswert als Männer. Hierin könnte sich zeigen, dass Frauen zu mehr Entspannung und Gelassenheit im ersten Behandlungsjahr als Männer gelangt sind, was sich in den anderen Skalen (MMPI[K] und GBB) nicht nachweisen lässt.

Die Unterschiede zwischen Männern und Frauen in den Skalenwerten von t3 nach t5 sind nur in der Skala Psychasthenie signifikant, Frauen gesunden - insoweit ihr Zwangsverhalten und ihre phobischen Gedanken im Fokus stehen - mehr als Männer.

3) Vergleichszeitraum t1 bis t5 (s. Tabelle 49 im Anhang):

In der reduzierten Grundgesamtheit der Patienten, von denen zu den Zeitpunkten t1 und t5 Daten vorliegen, sind die Unterschiede zwischen Männern und Frauen zum Zeitpunkt t1 nicht signifikant. Die Unterschiede zwischen t1 und t5 sind für die Gesamtgruppe in allen Skalen hoch signifikant, während die Unterschiede zwischen Männern und Frauen in der Entwicklung von t1 bis t5 in den vier MMPI(K)-Skalen signifikant sind, in den Skalen Depression und Hysterie sogar sehr signifikant. Für beide Skalen gilt, dass Frauen hier zum Zeitpunkt t1 ein höheres Ausgangsniveau hatten, zum Zeitpunkt t5 jedoch niedrigere Werte als die Männer. Dies gilt auch für die Skalen Hypochondrie und Psychasthenie. Depressive (MMPI[K]-D) und hysterische (MMPI[K]-HY) Frauen konnten demnach von der dreijährigen Behandlung besonders stark profitieren, mehr als die Männer und auch noch mehr als die Frauen mit hypochondrischen (MMPI[K]-HS) und psychasthenischen (MMPI[K]-PT) Merkmalen.

In den GBB-Skalen ist diese Entwicklung nicht nachweisbar, wenngleich auch hier, in GBB 5 (Beschwerdedruck), die Frauen etwas mehr von der Behandlung zu profitieren scheinen. Insofern also die somatoformen Beschwerden im Vordergrund stehen, lässt sich eine unterschiedliche Entwicklung von Männern und Frauen in der dreijährigen Behandlungszeit nicht überzeugend nachweisen.

*Zusammenfassung:* Frauen entwickeln sich auf den gesamten dreijährigen Beobachtungszeitraum bezogen in den Persönlichkeitsskalen des MMPI(K) günstiger in Richtung Gesundheit als Männer. In den beiden Zeitspannen t1 – t3 u. t3 – t5 fällt dieser Unterschied weniger eindeutig aus.

#### **10.4.3 Varianzanalysen mit Patientenalter bis 30Jahre, 31-40, über 40 Jahre alt im Zeitverlauf von t1 bis t3, t3 bis t5, t1 bis t5 und den Mittelwerten der GBB-Skalen 1 bis 5, der MMPI(K)-Skalen HS, D, HY, PT und des VEV**

1) Vergleichszeitraum t1 bis t3 (s. Tabelle 50 im Anhang):

Die Unterschiede in den Altersgruppen zum Zeitpunkt t1 sind in GBB-Gliederschmerzen signifikant, hier haben die über 40Jährigen einen deutlich höheren Wert als die beiden jüngeren Altersgruppen, Gliederschmerzen sind demzufolge mehr das Problem der Älteren.

Die Verringerung der Werte bis zum Zeitpunkt t3 differiert nicht signifikant zwischen den Altersgruppen, obwohl sich hier doch gewisse Tendenzen zeigen: Die Altersgruppe der über 40Jährigen scheint sich im Vergleich zu den anderen Altersgruppen in fast allen Skalen des GBB und den vier Skalen der neurotischen Trias +PT des MMPI(K) günstiger zu entwickeln, insofern die Spanne vom Ausgangswert zu t1 bis

zum Wert des Messzeitpunktes t3 jeweils größer ist als in den jüngeren Altersgruppen (außer in GBB-Magenbeschwerden), was zumeist an den niedrigeren Mittelwerten zu t3 liegt.

## 2) Vergleichszeitraum t3 bis t5 (s. Tabelle 51 im Anhang):

In der reduzierten Grundgesamtheit der Patienten, bei denen zu t3 und t5 Daten vorliegen, zeigen sich in den drei Altersgruppen zu t3 ähnliche Mittelwerte, in der Skala Hysterie des MMPI(K) fällt der niedrigere Wert der Altersgruppe „über 40“ auf, im VEV der 31 bis 40Jährigen, der Unterschied zu den anderen Altersgruppen ist jedoch nicht signifikant. Auch in GBB-Gliederschmerzen unterscheidet sich die Gruppe der über 40Jährigen von den beiden jüngeren Altersgruppen - hier durch einen höheren Mittelwert - aber auch dieser Unterschied ist nicht signifikant. Ein altersunabhängiger Zeiteffekt von t3 bis t5 mit unterschiedlicher statistischer Signifikanz zeigt sich bei GBB-Magenbeschwerden, GBB-Herzbeschwerden und GBB-Beschwerdedruck und in den MMPI(K)-Skalen, die altersbedingten Unterschiede von t3 bis t5 sind dagegen allesamt nicht signifikant, jedoch ergeben sich interessante Hinweise: In der GBB-Skala Beschwerdedruck erreichen die über 40Jährigen den relativ geringsten Mittelwert zu t5, obwohl sie zu t3 noch den höchsten Mittelwert der drei Altersgruppen aufwiesen. Auch in den Skalen GBB-Magenbeschwerden und GBB-Gliederschmerzen verringert sich der Mittelwert in dieser Altersgruppe deutlich und erreicht bzw. unterschreitet das Niveau der anderen Altersgruppen.

Auch für die MMPI(K)-Skalen gilt, dass die Altersgruppen sich unterschiedlich entwickeln, obwohl - außer bei Hysterie - unterhalb der statistischen Signifikanzgrenze. Auf der Skala Hysterie erreicht die Gruppe der bis 30Jährigen zu t5 einen höheren Mittelwert als zu t3, während die Gruppe der 31 bis 40Jährigen und die Gruppe der über 40Jährigen niedrigere Mittelwerte aufweisen. Dieselbe Relation gilt auch für die anderen MMPI(K)-Skalen, auch hier stagnieren die Werte der bis 30Jährigen, während die Werte der beiden älteren Gruppen abnehmen.

In den Mittelwerten des VEV lassen sich aus den Zahlen zu t3 und t5 interessante Hinweise auf Stagnation der mittleren Altersgruppe und „gesunde Weiterentwicklung“ der jüngeren, insbesondere aber der älteren, finden.

Möglicherweise zeigt sich bei einer Gesamtwürdigung, dass bei den jüngeren Patienten die Langzeitbehandlung verglichen mit kürzerer Therapiedauer keinen gut nachweisbaren zusätzlichen Nutzen bringt, was sich bei der Gruppe der über 40Jährigen anders darstellt. Hier zeigen sich deutlicher Langzeiteffekte, die mit kürzerer Therapiedauer vielleicht nicht hätten erreicht werden können.

## 3) Vergleichszeitraum t1 bis t5 (s. Tabelle 52 im Anhang):

Der Drei-Jahresvergleich, beginnend zu t1, zeigt ebenfalls keine signifikanten Unterschiede zwischen den drei Altersgruppen, mit Ausnahme der evtl. gerade noch tendenziellen Signifikanz des Unterschiedes in Bezug auf GBB-Gliederschmerzen, der durch das höhere Ausgangsniveau der über 40Jährigen zu t1 zustande kommt, die dann zu t5 ungefähr das gleiche Niveau wie die anderen Altersgruppen erreichen. Auffallend ist auch hier - für diesen Drei-Jahreszeitraum - der therapeutische Gewinn der über 40Jährigen, die in allen Skalen, außer in GBB-Gliederschmerzen, bedingt durch den zu t1 höheren Ausgangswert, „gesündere Endwerte“ (zu t5) haben als die jüngeren Altersgruppen.

*Zusammenfassend spricht einiges dafür, dass die Altersgruppe der über 40-Jährigen von der dreijährigen Behandlungszeit am stärksten profitiert hat. Den jüngeren Patienten scheint die Langzeitbehandlung nicht unbedingt einen Mehrnutzen gegenüber kürzeren Verläufen zu bringen.*

#### **10.4.4 Varianzanalysen mit den Therapeut-Patient-Dyaden Therapeut/Patient, Therapeutin/Patient, Therapeutin/Patientin, Therapeut/Patientin im Zeitverlauf von t1 bis t3, t3 bis t5, t1 bis t5 und den Mittelwerten der GBB-Skalen 1 bis 5, der MMPI(K)-Skalen HS, D, HY, PT und des VEV**

##### **1) Vergleichszeitraum t1 bis t3 (s. Tabelle 53 im Anhang):**

Die Bedeutung des Zusammenpassens von Patient und Analytiker als ein Prädiktor für die Langzeitwirkung konnte schon in der Boston Psychoanalytic Institute Prediction Study (Bachrach, 1991, s. Kap.2) nachgewiesen werden und ist auch später in vielen anderen Studien betont worden (so z.B. Orlinsky u. Howard, 1986, Hutterer, 1996).

Zum Zeitpunkt t1 unterscheiden sich die Subgruppen mit unterschiedlicher Signifikanzstärke. Zunächst fällt auf, dass die Dyade männlicher Therapeut-Patientin in nahezu allen Skalen (außer GBB-Herzbeschwerden) die höchsten Mittelwerte hat, in den vier Skalen der neurotischen Trias +PT des MMPI(K) sind die Unterschiede zwischen den vier Subgruppen signifikant bis tendenziell signifikant; signifikant bzw. tendenziell signifikant sind auch die Unterschiede in GBB-Erschöpfungsneigung, -Gliederschmerzen und -Beschwerdedruck. Von Therapiebeginn an bis zum Ende des ersten Behandlungsjahres zeigt sich in den Skalen nur ein (tendenziell) signifikanter Unterschied zwischen den verschiedenen Dyaden, und zwar in der Skala GBB-Magenbeschwerden. Hier machen Therapeut/Patientin die größten Fortschritte, während Therapeutin/Patient nur eine geringe Veränderung hervorbringen, dies gilt auch für GBB-Beschwerdedruck. Das durchweg signifikante p im Cochran-Test, insbesondere zu t3, zeigt die Heterogenität der Gruppen an, was allerdings die Interpretation der Varianzanalyse in Frage stellt.

##### **2) Vergleichszeitraum t3 bis t5 (s. Tabelle 54 im Anhang):**

Auf der Basis der Patienten, von denen zum Zeitpunkt t3 und zu t5 Daten vorliegen, sind die Unterschiede zwischen den Dyaden zum Zeitpunkt t3 in GBB-Erschöpfungsneigung, -Gliederschmerzen, -Beschwerdedruck und in MMPI(K)-Psychasthenia tendenziell bis sehr signifikant, jedoch schränkt das niedrige N=7 in der Dyade Therapeutin/Patient die Zuverlässigkeit dieses Befundes ein, das gleiche gilt auch für die Signifikanz des Unterschieds im Zeitverlauf in GBB-Magenbeschwerden und GBB-Gliederschmerzen, sodass wir auf eine Interpretation verzichten. Vielleicht sind einige Tendenzen erwähnenswert: Die Dyade Therapeut/Patientin ist auffallend erfolgreich in den vier MMPI(K)-Skalen, die Dyade Therapeutin/Patientin in GBB-Herzbeschwerden und ebenfalls in den MMPI(K)-Skalen, besonders Psy-

chasthenie, die Dyade Therapeut/Patient in GBB-Erschöpfungsneigung, die Dyade Therapeutin/Patient zeigt in fast allen Skalen positive Veränderungen, jedoch, basierend auf einem N=7, kaum interpretierbar. Geringste Wirkung (Zeitraum zweites bis drittes Therapiejahr) erzielt in den Skalen des MMPI(K) die Dyade Therapeut/Patient, in der sich ca. 20 % aller Patienten dieses Zeitraumes befinden. Alle anderen Paarungen scheinen in mehr oder weniger großem Ausmaß effektiver zu sein, soweit dies in den MMPI(K)-Skalen sichtbar wird.

Im VEV erreicht die Dyade Therapeut/Patientin zum Beobachtungsende nach dreijähriger Behandlung den höchsten Veränderungswert.

### 3) Vergleichszeitraum t1 bis t5 (s. Tabelle 55 im Anhang):

Die durch das niedrige N erwähnten Einschränkungen gelten für diese Varianzanalyse entsprechend. Auf der Basis der Patienten, bei denen zu t1 und zu t5 Daten vorliegen, ergibt sich zum Zeitpunkt t1 ein tendenziell signifikanter Unterschied zwischen den Dyaden bei GBB-Erschöpfung und ein signifikanter Unterschied in den MMPI(K)-Skalen Psychasthenia und Hysterie. In allen drei Skalen sind die Mittelwerte für die Dyade Therapeut/Patientin höher als für die anderen Dyaden, in den anderen Skalen ist dies ebenfalls der Fall, ohne dass hier statistisch signifikante Unterschiede zu verzeichnen wären.

Vom Verlauf her zeigt sich ein sehr signifikanter Unterschied in der MMPI(K)-Skala Depression und ein signifikanter Unterschied in der MMPI(K)-Skala Hysterie. Während die männlichen Patienten, egal ob sie jetzt bei einem männlichen Therapeuten oder bei einer Therapeutin in Therapie sind, nur relativ geringe Veränderungen aufweisen, verändern sich die Frauen in den Skalen Depression und Hysterie (auch in Psychasthenie und Hypochondriasis) mehr, und zwar bei männlichen und weiblichen Therapeuten nahezu gleich.

*Zusammenfassung:* Zu den männlichen Therapeuten gehen offensichtlich die stärker beeinträchtigten Frauen in Behandlung, die dann bis zum Ende des dritten Behandlungsjahres in den Neuroseindikatoren des MMPI(K) den höchsten Rückgang erreichen und zur relativ größten Entspannung und Gelassenheit gelangen. Auch die Behandlung von Magenbeschwerden zeigt die größten Fortschritte in der Dyade Therapeut/Patientin, während Therapeutin/Patient dort und im Beschwerdedruck nur eine geringe Veränderung hervorbringen.

Während die männlichen Patienten im gesamten Behandlungszeitraum (unabhängig vom Therapeutengeschlecht) nur relativ geringe Veränderungen aufweisen, verändern sich die Frauen in den Neuroseindikatoren des MMPI(K) stärker, und zwar bei männlichen und weiblichen Therapeuten nahezu gleich.

#### **10.4.5 Varianzanalysen mit Sitzungsfrequenz (niedriger/höher), den Messzeitpunkten nach einem (t3), nach zwei (t4) und nach drei (t5) Jahren Behandlung, im Zeitverlauf von t1 bis t3, t3 bis t4, t3 bis t5, t4 bis t5 und den Mittelwerten der GBB-Skalen 1 bis 5, der MMPI(K)-Skalen HS,D, HY, PT und des VEV**

In diesem Kapitel wenden wir uns der Frage zu, welchen Einfluss die wöchentliche Stundenfrequenz auf den Therapieeffekt hat. Nach Hartkamp (1997) ist dieser Zusammenhang in der psychoanalytischen Psychotherapie noch weitgehend ungeklärt und wird deshalb in der zukünftigen Forschung eine wichtige Rolle spielen (s. Kap. 13). Wir wollten wissen, ob die Stundenfrequenz im Verlauf der Behandlung die (Zwischen-)Ergebnisse so beeinflusst, dass ein statistisch bedeutsamer Unterschied zwischen niedrigerer und höherer Sitzungsanzahl nachweisbar ist. Zu diesem Zweck haben wir drei Messzeitpunkte festgelegt: nach einem (t3), nach zwei (t4) und nach drei Jahren (t5) Behandlung. Für die bis zu diesen Zeitpunkten jeweils stattgefundenen Sitzungen pro Patient wurde der Median gebildet. Die Teilgruppe unterhalb des Median nennen wir die „Niederfrequenten“, die oberhalb des Median die „Höherfrequenten“.

Die falsifizierbare Ho-Forschungshypothese lautet: Die Frequenz (niedriger/höher) der Stunden bis zu den drei Messzeitpunkten t3, t4, t5 hat keine Auswirkung auf die Messergebnisse in den fünf GBB-, den MMPI(K)-Skalen HS, D, HY, PT und dem VEV zu diesen drei Zeitpunkten, gemessen von dem jeweils ein Jahr vorher liegenden Zeitpunkt aus.

In Tabelle 56 sind die Sitzungsmediane und die Größe der Grundgesamtheiten zu t3, t4 und zu t5 aufgeführt.

**Tabelle 56: Anzahl der Sitzungen - Mediane -  
auf der Basis der Grundgesamtheiten  
zu t3, t4 und t5 in der UG**

	<b>Median/ Sitzungen</b>	<b>N</b>
<b>t3</b>	70	106
<b>t4</b>	137	79
<b>t5</b>	198	72

Aus den Sitzungsmedianen lässt sich die mittlere wöchentliche Stundenfrequenz berechnen. Im ersten Behandlungsjahr betrug sie 1,56 Sitzungen, bis zum Ende des zweiten Jahres 1,52 Sitzungen, bis zum Abschluss des Untersuchungszeitraumes nach drei Jahren 1,47 Sitzungen, berechnet jeweils unter der Annahme von sieben Wochen therapiefreie Zeit (Urlaub, Feiertage usw.) pro Jahr.

Die durchschnittliche Stundenfrequenz nimmt also bis zum Ende des dritten Jahres geringfügig ab.

## Ergebnisse der Varianzanalysen:

### 1) Vergleichszeitraum t1 bis t3 (s. Tabelle 57 im Anhang):

Zum Zeitpunkt t1 weisen die Subgruppen keinen signifikanten Unterschied auf. Die Unterschiede zwischen den beiden Gruppen der Niedrigerfrequenten und der Höherfrequenten (Trennung über dem Median zu t3) von t1 bis t3 sind in allen untersuchten Variablen nicht signifikant. Das heißt, die festgestellte Wirksamkeit ist unabhängig von der Stundenanzahl nachweisbar. Ob nieder- oder höherfrequent spielt in der Zeit von t1 bis t3, dem ersten Behandlungsjahr, für das Ausmaß der Veränderung keine Rolle.

### 2) Vergleichszeitraum t3 bis t4 (s. Tabelle 58 im Anhang):

Zum Zeitpunkt t3 weisen die Subgruppen, wie auch schon zu t1, keinen signifikanten Unterschied auf. In der Gesamtgruppe lässt sich ein signifikanter Zeiteffekt von t3 nach t4 in den Skalen Hypochondrie und Hysterie des MMPI(K) nachweisen.

Der Vergleich der durch Teilung über dem Median zu t4 entstandenen Gruppe mit eher weniger Stunden (niedrigerfrequent) mit der Gruppe mit eher mehr Stunden (höherfrequent) ergibt sehr signifikante Unterschiede (GBB-Gliederschmerzen, GBB-Beschwerdedruck), und signifikante Unterschiede (GBB-Erschöpfungsneigung und -Herzbeschwerden) und tendenziell signifikante Unterschiede (Hypochondrie) im Zeitverlauf von t3 bis t4. Während bei der Gruppe mit niedrigerfrequenten Sitzungen in diesen Skalen sogar die Mittelwerte leicht ansteigen, gehen sie bei der Gruppe mit höherfrequenten Stunden zurück. In MMPI(K)-Hypochondrie fallen die Mittelwerte von t3 nach t4 sowohl in der niedrigerfrequenten Gruppe als auch in der höherfrequenten Gruppe, jedoch in der höherfrequenten Gruppe deutlich ausgeprägter.

Es zeigt sich gerade in diesem Zeitabschnitt der Behandlung, dass für eine weitere Verbesserung - nach der positiven Entwicklung der psychosomatischen Symptome im ersten Behandlungsjahr - eine niedrigfrequente Behandlung nicht mehr ausreicht, sondern eine höhere Sitzungsanzahl notwendig ist.

### 3) Vergleichszeitraum t4 bis t5 (s. Tabelle 59 im Anhang):

Nachdem die gesamte Untersuchungsgruppe zum Zeitpunkt t5 wiederum über dem Median der Sitzungsanzahl in zwei gleichgroße Gruppen, „Sitzungen eher niedrigerfrequent“ und „Sitzungen eher höherfrequent“ geteilt worden ist, zeigt sich zum Zeitpunkt t4 kein signifikanter Unterschied in den untersuchten Skalen und im VEV. In MMPI(K)-Depression, -Hysterie und -Psychasthenie ist der jeweilige Mittelwert in der Gruppe der Niedrigerfrequenten niedriger, im VEV höher als in der Gruppe der Höherfrequenten. Dies hängt möglicherweise mit einer Reduktion der Stundenanzahl bei den im Gesundungsprozess Fortgeschritteneren und einer damit verbundenen stärkeren Konzentration der Kränkeren in der Gruppe der Höherfrequenten zusammen. Für die Gesamtgruppe lässt sich ein sehr signifikanter Zeiteffekt bei MMPI(K)-Psychasthenie und ein knapp signifikanter Zeiteffekt bei GBB-Herzbeschwerden und bei MMPI(K)-Depression feststellen.

Die Unterschiede zwischen den beiden Gruppen der Niedrigerfrequenten und Höherfrequenten im Zeitverlauf von t4 nach t5 sind allesamt nicht signifikant.

Dies kann bedeuten, dass in diesem Zeitraum, drittes Behandlungsjahr, wie auch im ersten Jahr, die Entwicklung der Patienten nicht von der Frequenz der Sitzungen pro Woche abhängt.

#### 4) Vergleichszeitraum t3 bis t5 (s. Tabelle 60 im Anhang):

Die durch die Trennung über dem Median zum Zeitpunkt t5 entstandenen beiden Frequenzgruppen (Niedrigfrequente/Höherfrequente) unterscheiden sich zum Zeitpunkt t3 tendenziell signifikant in den Mittelwerten von GBB-Gliederschmerzen, MMPI(K)-Depression und MMPI(K)-Hysterie. Die Mittelwerte in der Gruppe der Höherfrequenten sind in den GBB- und MMPI(K)-Skalen durchweg höher als in der Gruppe der Niedrigfrequenten, mit Ausnahme für GBB 2 (Magenbeschwerden), hier ist zu t3 der Mittelwert bei den Höherfrequenten niedriger als bei den Niedrigfrequenten.

Bei Betrachtung des Zweijahreszeitraumes von t3 bis t5 konnten die Patienten mit Magenbeschwerden von der Therapie mehr profitieren, wenn die Behandlung höherfrequent durchgeführt wurde, in allen anderen Skalen von GBB und MMPI(K) sind statistisch signifikante frequenzabhängige Unterschiede nicht nachweisbar. Die von t3 bis t4 festgestellten Effekte (s.o.) sind in der Zweijahresbetrachtung (mit reduzierter Grundgesamtheit) nicht auffindbar.

Auch im VEV sind die Mittelwerte zwischen den beiden Frequenzgruppen zu t3 unterschiedlich. Die Gruppe mit höherer Stundenfrequenz zu t5 schätzt die bis zu t3 erfolgte Veränderung im Erleben und Verhalten geringer ein als die Gruppe mit geringerer Stundenfrequenz. Während die Patienten mit geringerer Stundenfrequenz sich dann von t3 bis t5 im Erleben und Verhalten kaum noch verändern (eher sogar wieder niedrigere Werte haben), können sich die Patienten mit höherer Stundenfrequenz positiv entwickeln, die Unterschiede sind gerade noch tendenziell signifikant.

#### *Zusammenfassung:*

Die nachweisbare Wirksamkeit der Behandlung im ersten Behandlungsjahr ist nicht an die Frequenz gebunden.

Im zweiten Behandlungsjahr ist für eine weitere Verbesserung der psychosomatischen Symptome - nach der positiven Entwicklung im ersten Behandlungsjahr - ein höherfrequentes Vorgehen nötig.

Im dritten Behandlungsjahr sind frequenzabhängige Unterschiede in der Wirksamkeit nicht feststellbar. Wie an anderer Stelle gezeigt, stagniert die Entwicklung zu mehr Gesundheit in der Gesamtgruppe der Patienten, lässt sich statistisch abgesichert aber noch an der MMPI(K)-Skala Psychasthenie und für Herzbeschwerden nachweisen.

Die Zweijahresbetrachtung (t3-t5) zeigt, dass die höhere Frequenz bei Patienten mit Magenbeschwerden effektiver ist, auch scheint sie sich auf Entspannung und Gelassenheit günstiger als die niedrigere Frequenz auszuwirken.

Die eingangs des Kapitels formulierte Ho-Hypothese lässt sich weder eindeutig bestätigen noch zurückweisen. Frequenzabhängige Unterschiede in den Ergebnissen variieren offensichtlich mit den Zeitpunkten während der Behandlung und den Symptomen bzw. den Messverfahren. Unsere Befunde zeigen eine Überlegenheit des höherfrequenten Vorgehens in Bezug auf psychosomatische Beschwerden in der Mitte des Beobachtungszeitraumes (zweites Behandlungsjahr).



#### 10.4.6 Varianzanalyse mit der Variablen „Therapeutengeschlecht“ (vgl. Tabellen 61 – 63 im Anhang)

Unsere theoriegeleitete Anfangsfrage lautete, ob ein unterschiedliches Therapeutengeschlecht unterschiedliche Auswirkungen auf den Erfolg der Therapie hat oder nicht. Hierbei ging es uns nicht um eine individuelle Einzelfallbetrachtung, bei der die Frage des Therapeutengeschlechts natürlich immer eine wichtige Rolle spielt (Matchingprozess), sondern hier ging es uns einzig und allein um eine statistische Sicht, nämlich ob in unserem Datenmaterial Tendenzen dahingehend ablesbar sind, dass es einen statistischen Zusammenhang zwischen der Wirksamkeit von Therapien und dem Geschlecht der Therapeuten gibt. Dabei verhehlen wir nicht unsere aus geschlechtsharmonischen Gründen gehegte Hoffnung, es mögen sich hoffentlich keine Wirksamkeitsunterschiede zwischen den Therapeutengeschlechtern nachweisen lassen. Allerdings liegt zu dieser Frage bereits ein anderslautendes Ergebnis aus dem Stockholmer Projekt vor: Dort hatten „Patienten bei Therapeutinnen ... signifikant bessere Ergebnisse als Patienten bei männlichen Therapeuten, unabhängig vom Geschlecht des Patienten und der Art der Behandlung“. (Sandell, 2001, S. 295)

Von unserer Anfangsfrage geleitet ließ sich die formal korrekt falsifizierbare Ho-Forschungshypothese wie folgt formulieren: ‚Es besteht kein nachweisbarer Zusammenhang zwischen Therapeutengeschlecht und Therapieerfolg.‘ Um dieser Hypothese nachzugehen, bildeten wir zunächst zwei statistische Untersuchungsgruppen wie folgt: Die Verschickung der Fragebögen an die Patienten hatte sich über die mitwirkenden Kollegen vollzogen, die das Testmaterial vor der Weiterleitung an die Patienten mit ihrem Namen und der Chiffre des Patienten versehen hatten. Von daher war es nun leicht möglich, die Patienten nach dem Geschlecht der sie behandelnden Therapeuten in zwei Subgruppen einzuteilen und eine Varianzanalyse durchzuführen.

Ein zunächst durchgeführter Mittelwertvergleich zeigte, dass sich im Querschnittvergleich der Subgruppen weder im GBB noch im MMPI(K) zu irgend einem der verglichenen Messzeitpunkte t1, t3 und t5 ein signifikanter Unterschied zwischen den beiden Subgruppen zeigte. Dies spricht dafür, dass die Patienten sich eher zufällig auf die beiden Therapeutengeschlechter verteilt hatten und die Wahl eines männlichen oder weiblichen Therapeuten kein Indikator für das Vorliegen einer bestimmten Art von Störung zu sein scheint. Entsprechend ist es auch nicht sehr überraschend, dass sich in der Abfolge von Messzeitpunkt t1 zu t3 und von t3 zu t5 kein signifikanter Verlaufsunterschied zwischen den beiden Subgruppen herausstellte. Dies bedeutet nicht, dass innerhalb der Subgruppen kein Therapieerfolg nachweisbar war, - sondern lediglich, dass sich die beiden Subgruppen im Ausmaß des Therapieerfolgs nicht signifikant voneinander unterschieden.

Im VEV zeigten sich in der längsschnitorientierten Auswertung der Messzeitpunkte t3 und t5 leicht unterschiedliche Verläufe der beiden Subgruppen: Zum Zeitpunkt t3 unterscheiden sich die beiden Subgruppen lediglich durch einen einzigen Mittelwertpunkt voneinander, was heißt, dass beide Subgruppen die durch die Therapie erzielte Veränderung im Therapiezeitraum von t1 zu t3 in etwa gleich groß einschätzen. Zum Messzeitpunkt t5, also in der Bewertung der eingetretenen Veränderungen zwischen t3 und t5, steigen der Mittelwert der Subgruppe „weiblicher Therapeut“ nur unwesentlich um drei Punkte, der Mittelwert der Subgruppe „männlicher Therapeut“ dagegen um zwölf Punkte an. Das heißt, dass die Patien-

ten bei männlichen Therapeuten die Veränderungen im Beobachtungszeitraum etwas höher einschätzen als die Patientengruppe bei weiblichen Therapeuten.

Dieser Unterschied erreicht jedoch kein Signifikanzniveau: Weder unterscheiden sich die Subgruppen zum Zeitpunkt der Messung signifikant voneinander, noch zeigen sie hinsichtlich ihrer Therapieeinschätzung im Zeitverlauf eine signifikant unterschiedliche Entwicklung, noch zeigt sich ein Gesamt-Psychotherapieeffekt dahingehend, dass sich die „Veränderungseinschätzung der Gesamtgruppe“, die um acht Mittelwertpunkte ansteigt, zwischen den Messzeitpunkten im Ganzen signifikant verändert.

Als Gesamtergebnis bleibt festzuhalten, dass die Ho-Hypothese bestätigt werden konnte: Es lässt sich zumindest in unseren Daten kein statistischer Zusammenhang zwischen einem bestimmten Therapeutengeschlecht und dem Therapieerfolg nachweisen. Die Therapien der beiden Therapeutengeschlechter sind in etwa gleich erfolgreich verlaufen.

#### **10.4.7 Varianzanalyse mit der Variablen „Therapeutenberuf“ (vgl. Tabellen 64 – 66 im Anhang)**

Unsere theoriegeleitete Anfangsfrage lautete: Sind die Psychotherapien der ärztlichen Psychotherapeuten wirkungsvoller als die der psychologischen Psychotherapeuten, oder verhält es sich umgekehrt, oder gibt es vielleicht gar keine statistisch relevanten Unterschiede? Auch hier müssen wir zugestehen, dass wir, - diesmal aus berufspolitischen Gründen -, hofften, dass der Erfolg von Psychotherapien, - unabhängig von der jeweils im Einzelfall relevanten persönlichen Qualifikation des Therapeuten -, so doch im statistischen Sinne eher nichts mit der Zugehörigkeit zu einem Grundberuf zu tun hat.

Unsere falsifizierbare Ho-Forschungshypothese lautete: ‚Es besteht kein nachweisbarer Zusammenhang zwischen dem Erfolg einer Therapie und dem Grundberuf des Psychotherapeuten.‘ Zur Untersuchung dieser Frage bildeten wir wieder zwei Untersuchungsgruppen: Mit dem Erstfragebogen zum Messzeitpunkt t1 hatten wir den „Therapeutenberuf“ erhoben, so dass wir jetzt ohne Schwierigkeiten in der Lage waren, alle unsere Patienten auf die statistischen Subgruppen „ärztlicher Psychotherapeut“ oder „psychologischer Therapeut“ zu verteilen. Dabei stellte sich ein zahlenmäßiges Ungleichgewicht heraus, dass nämlich von 104 auswertbaren Patienten zum Zeitpunkt t3 nur 14 Patienten bei einem ärztlichen Psychotherapeuten, dagegen 90 Patienten bei einem psychologischen Psychotherapeuten in Behandlung waren. Dieses Ungleichgewicht hat sich von t1 über t3 bis zu t5 fortgesetzt, - allerdings glücklicherweise so, dass die Anzahl (N) der Patienten bei den ärztlichen Kollegen weitgehend konstant blieb und das geringer werdende Gesamt-„N“ sich hauptsächlich durch einen Patientenschwund bei den psychologischen Psychotherapeuten ergab, - so dass sich die Subgruppe der bei ärztlichen Therapeuten in Behandlung befindlichen Patienten nicht auflöste und damit die Auswertung (mit allen Vorbehalten wegen des doch insgesamt sehr geringen „N“) weiter möglich war.

Unsere Varianzanalyse kam zu folgenden Ergebnissen: Zunächst einmal mussten wir feststellen, dass sich die Patienten, gleich ob sie bei einem ärztlichen oder psychologischen Psychotherapeuten in Behandlung waren, in ihrem Antwortverhalten auf den klinischen Skalen des GBB und des MMPI(K) mit Ausnahme von drei Skalen (Validity, Psychopathie und Paranoia) zum Messzeitpunkt t1 nicht signifikant voneinander unterschieden. Gleichwohl zeigten sich unterhalb des Signifikanzniveaus bei der Patientengruppe der psychologischen Psychotherapeuten zum Messzeitpunkt t1 leicht höhere Mittelwertausprä-

gungen in den Skalen Magenbeschwerden, Gliederschmerzen, Lie und Correction. In allen anderen Skalen zeigten die bei den Ärzten in Behandlung befindlichen Patienten leicht höhere Ausprägungen.

Die Abweichungen bleiben insgesamt jedoch so niedrig, dass wir von einer zufallsbedingten Erscheinung ausgehen. Was die Gesamtheit der Patientenpopulation ohne Berücksichtigung der Aufteilung in Subgruppen anbelangt, so konnten wir in dem Vergleich der Messzeitpunkte t1 zu t3 feststellen, dass in acht klinischen Skalen des GBB und MMPI(K) eine ‚hochsignifikante‘ Veränderung im Sinne eines positiven Therapieeffekts und in vier weiteren Skalen eine ‚sehr signifikante‘ Wirkung nachzuweisen ist. Diese Tendenz setzt sich von t3 zu t5 in drei ‚sehr signifikant‘ und fünf noch ‚signifikant‘ anzeigenden Skalen fort und bildet sich im überlagernden Gesamtvergleich von t1 zu t5 erwartungsgemäß noch deutlicher ab. Dieser Therapieerfolg im Sinne einer Symptomminderung scheint sich jedoch auf beide Subgruppen gleich zu verteilen, - zumindest zeigt sich kein signifikanter Verlaufsunterschied im Zeitvergleich der beiden Subgruppen.

Im VEV-Vergleich t3 zu t5 sind die Unterschiede der Mittelwerte zwischen der Subgruppe „ärztlicher Psychotherapeut“ und „psychologischer Psychotherapeut“ zwar nicht sehr hoch (30 Punkte), erreichen aber gleichwohl mit dem Wert  $F = 0.021$  noch das niedrigste Signifikanzniveau. Die beiden Subgruppen unterscheiden sich insofern, dass die Patienten der psychologischen Psychotherapeuten mit 210 Punkten eine tendenziell bessere Therapieeinschätzung abgeben als die Patientengruppe bei den ärztlichen Psychotherapeuten, die eine mittlere Veränderungseinschätzung in der Größenordnung von 180 Punkten abgeben. Zum Messzeitpunkt t5, also in der Einschätzung der eingetretenen Veränderungen zwischen t3 und t5, erhöhen sich die Mittelwerte der Subgruppe „ärztlicher Psychotherapeut“ nur unwesentlich um zwei Punkte, der Mittelwert der Subgruppe „psychologischer Psychotherapeut“ bleibt unverändert bei dem Wert von 210 Punkten. Das heißt, dass sich die beiden Subgruppen hinsichtlich ihrer „Veränderungseinschätzung durch Psychotherapie“ zwischen den beiden Messzeitpunkten nicht bzw. so gut wie nicht verändert haben. Die Subgruppen bleiben tendenziell signifikant unterschiedlich in ihrer Einschätzung, zeigen in dieser Unterschiedlichkeit im Zeitverlauf jedoch weder eine signifikant unterschiedliche Entwicklung, noch zeigt sich ein Gesamt-Psychotherapieeffekt dahingehend, dass sich die Veränderungseinschätzung der Gesamtgruppe zwischen den Messzeitpunkten im Ganzen signifikant verändert.

Unsere Ho-Hypothese konnte also bestätigt werden: Es besteht kein signifikanter Zusammenhang zwischen dem Erfolg einer Psychotherapie und dem Grundberuf des durchführenden Psychotherapeuten. Beide Berufsgruppen haben im statistischen Sinne „objektiv“ vergleichbar gute Therapieerfolge. Allerdings werden diese „objektiv“ gleichen Therapieerfolge von den Patienten „subjektiv“ tendenziell signifikant unterschiedlich wahrgenommen (vgl. VEV). Bei einer in unserer Untersuchung objektiv nachgewiesenen gleichen Therapiewirksamkeit der beiden Grundberufe schätzen die von „Psychologen“ behandelten Patienten ihren Therapieerfolg größer ein als die von „Ärzten“ behandelten Patienten, wobei sich diese Einschätzung von Messzeitpunkt t3 zu t5 nicht verändert, sondern (bis auf eine marginale Abweichung) überraschenderweise sehr konstant bleibt. Es wäre nun ein wissenschaftslogischer Fehler, die „Einschätzung“ der Patienten hinsichtlich der Therapieerfolge mit den „tatsächlichen Therapieerfolgen“ zu verwechseln. An dieser Stelle würden denn auch zurecht wissenschaftslogische Zweifel der empirischen Sozialforschung an solchen katamnestischen Untersuchungen ansetzen, die sich lediglich auf das Datenmaterial der retrospektiven „Veränderungseinschätzungen“ der Patienten, nicht aber auf Prä-Post-Untersuchungen stützen. Wir möchten hier noch einmal an Langenmayr und Kosfelder (1998, S. 103)

erinnern, die (wie oben bereits ausführlich diskutiert) auf den Unterschied zwischen aktuell gemessenen klinischen Daten und retrospektiven Einschätzungen klinischer Beschwerden bzw. deren Veränderung aufmerksam machen.

#### **10.4.8 Varianzanalyse mit der Variablen „Schulabschluss“ (vgl. Tabellen 67 – 69 im Anhang)**

Unsere theoriegeleitete Anfangsvermutung lautete, dass die Sozialisation des Patienten, definiert über die Art des Schulabschlusses, mit der Wirkung von Psychotherapie derart korreliert, dass Patienten mit höherem Schulabschluss prinzipiell eher von Psychotherapie profitieren als Patienten mit niedrigerem Schulabschluss. „Therapieerfolg“ wird hier wieder durch den Unterschied in der Ausprägung auf den klinischen Skalen des GBB, MMPI(K) und VEV zu den Messzeitpunkten t1, t3 und t5 operationalisiert. Daraus ließ sich unsere falsifizierbare Ho-Forschungshypothese wie folgt ableiten: ‚Zwischen der Art des Schulabschlusses einer Patientengruppe und dem Ausmaß der erzielbaren positiven Wirkung von Psychotherapie besteht kein signifikanter statistischer Zusammenhang.‘

Die Bildung der statistischen und zu vergleichenden Untersuchungsgruppen vollzog sich wie folgt: Wir teilten die Patienten in drei Gruppen ein, nämlich in die Gruppe mit dem Schulabschluss „Hauptschule oder schlechter“, die Gruppe mit der „Mittleren Reife“ und vergleichbarer Schulabschlüsse sowie die Gruppe mit dem Abschluss „Abitur und ggf. auch Studium“.

Zunächst konnten wir feststellen, dass sich die so gebildeten drei Subgruppen im Mittelwertvergleich ihrer klinischen Skalen weder zu t1, noch zu t3 und auch nicht zu t5 signifikant unterscheiden, was heißt, dass alle drei Subgruppen zu allen drei Messzeitpunkten eine gleich umfängliche klinische Symptombelastung aufwiesen. Von daher scheint es dann allerdings keine Überraschung mehr zu sein, dass sich auch hinsichtlich des Verlaufs der Subgruppen von t1 zu t3 und dann von t3 zu t5 kein signifikanter Unterschied zeigte. Dass die zusammengefasste Großgruppe aller Patienten, also unabhängig vom Schulabschluss gesehen, in den klinischen Skalen wieder ‚hoch‘ und ‚sehr signifikante‘ Mittelwertsenkungen im Sinne von Symptomminderungen ausweist, entspricht nach den bisherigen Ergebnissen unserer Untersuchung der Erwartung, da es sich ja unabhängig von der Aufteilung in unterschiedlich kriterienbezogene Subgruppen ja jedes mal um die gleiche Gesamtpatientengruppe handelt.

Auch im VEV-Vergleich t3 zu t5 sind die Unterschiede der Mittelwerte zwischen den Subgruppen „Hauptschule“, „Mittlere Reife“ und „Abitur“ sehr gering und erreichen kein Signifikanzniveau. Die Subgruppen unterscheiden sich also fast gar nicht (maximaler Unterschied von vier Mittelwertpunkten) voneinander. Alle drei Subgruppen schätzen die durch die Therapie erzielte Veränderung im Therapiezeitraum t1 zu t3 in etwa gleich groß ein. Zum Messzeitpunkt t5, also in der Bewertung der eingetretenen Veränderungen zwischen t3 und t5, steigen die Mittelwerte der Subgruppe „Hauptschule“ und „Abitur“ jeweils um zwölf Punkte, der Mittelwert der Subgruppe „Mittlere Reife“ dagegen fällt um neun Punkte. Das heißt, dass die Subgruppen „Hauptschule“ und „Abitur“ die Veränderungen im Beobachtungszeitraum positiv einschätzen, während die Subgruppe „Mittlere Reife“ eine negative Bewertung abgibt. Keiner dieser Werte erreicht jedoch ein Signifikanzniveau: weder gibt es zum Zeitpunkt der Messung einen signifikanten Unterschied zwischen den Subgruppen, noch zeigen sie im Zeitverlauf eine signifikant unterschiedliche Entwicklung der Therapieeinschätzung, - noch zeigt sich ein Gesamt-Psychotherapieeffekt dahingehend,

dass sich die „Veränderungseinschätzung der Gesamtgruppe“ zwischen den Messzeitpunkten im Ganzen signifikant verändert.

Unsere Ho-Hypothese, wonach kein nachweisbarer Zusammenhang zwischen der Art des Schulabschlusses und dem Ausmaß der Wirksamkeit von Psychotherapie besteht, konnte also bestätigt werden. Alle Patientengruppen, gleich mit welchem Schulabschluss, haben umfänglich vergleichbar hochsignifikante Therapieergebnisse. Dies entspricht nicht ganz der landläufigen Meinung, dass das Ausmaß der Wirksamkeit von Psychotherapie von der Sozialisation, hier definiert über die Schulbildung, abhängt.

#### **10.4.9 Varianzanalyse mit der Variablen „Anfänglichem Lie-Verhalten“ (vgl. Tabellen 70 – 72 im Anhang)**

Man kann in der Lie-Skala des MMPI(K) das messtheoretische Äquivalent für das Ausmaß des psychischen Widerstands gegen therapeutische Veränderungen sehen. Unsere theoriegeleitete Anfangsvermutung lautete, dass Patienten mit einem hohen Lie-Wert, also mit einer hohen Auskunftsbwehrhaltung, eine eher ungünstige Therapieprognose haben, wohingegen Patienten mit einem eher niedrigeren Lie-Wert, also mit einer eher naiv-unbefangeneren Auskunftsbereitschaft, eine eher günstige Prognose haben.

Darüber hinaus kann man auch vermuten, dass die Patienten mit einem eher hohen Lie-Wert (hoher Widerstand) die tendenziell eher kränkere Patientengruppe im Vergleich zu der eher gesunderen Patientengruppe mit einem eher niedrigeren Lie-Wert (geringer Widerstand) darstellen, insofern nämlich das Ausmaß des Widerstandes mit dem Ausmaß der neurotischen Verfestigung einhergeht. Weiter kann man dann vermuten, dass die Patienten mit den höheren Lie-Werten gerade wegen ihrer Widerständigkeit zum Messzeitpunkt t1 höhere Mittelwerte in den klinischen Skalen erzielen müssten als die Patientengruppe mit den niedrigeren Lie-Werten. Als Therapieerfolg legten wir wieder das Ausmaß der objektiv messbaren Veränderungen in den bekannten klinischen Skalen unserer Messverfahren fest.

Die falsifizierbare Ho-Forschungshypothese lautete, dass es keinen signifikanten statistischen Zusammenhang zwischen dem unterschiedlichen Ausmaß an Lie-Verhalten (Widerstand) und dem Ausmaß von Psychotherapieerfolg gibt. Die Bildung unserer statistischen Vergleichs- und Untersuchungsgruppen „widerständigere und mit höheren Lie-Werten besetzte Gruppe“ und „therapiebereitere und mit niedrigeren Lie-Werten besetzte Gruppe“ erfolgte empirisch, und zwar über das Antwortverhalten der Patienten in der Lie-Skala des MMPI(K). Über das statistische Maß des Median trennten wir die Gruppe der Probanden/Patienten mit den höheren Lie-Werten von der Gruppe der Probanden/Patienten mit den niedrigeren Lie-Werten.

Die Mittelwertvergleiche zeigen, dass sich die Gruppen „geringes Anfangs-Lie“ und „hohes Anfangs-Lie“ in den Ausgangsmittelwerten der klinischen Skalen zu den Messzeitpunkten t1, t3 und t5 kaum voneinander unterscheiden, dass die beiden statistisch gebildeten Lie-Subgruppen auf den klinischen Skalen also nicht diskriminieren - mit Ausnahme der Lie-Skala selbst, - was aber kein Messergebnis sondern nur Ausdruck unserer künstlich hergestellten Gruppenbildung unter Zuhilfenahme der Ausprägungen in dieser Skala ist. Für die Gesamtgruppe der Patienten, also unter Weglassung der Lie-Subgruppeneinteilung, - und dies war ja abermals nicht anders zu erwarten -, zeigen sich wieder hochsignifikante positive Ver-

änderungen in der Mehrzahl der klinischen Skalen: zwischen den Messzeitpunkten t1 und t3 zeigen zwölf klinische Skalen ‚hochsignifikante‘ Veränderungen, zwischen t3 und t5 zwei Skalen ‚sehr signifikante‘ Veränderungen und vier klinische Skalen ‚signifikante‘ Veränderungen im Sinne von gesundheitlich günstiger Symptomminderung an. Die überlagerte Darstellung der Ergebnisse für den Gesamtzeitraum t1 bis t5 zeigt wieder die bereits bekannten hochsignifikanten Erfolge in fast allen Skalen. Eine besondere Entwicklung ist interessanterweise in der Lie-Skala des MMPI(K) zu beobachten. Diese reagiert geradezu atypisch: der Mittelwert der Gruppe mit geringem Anfangs-Lie steigt von 1,914 zum Zeitpunkt t1 auf 2,882 zum Zeitpunkt t5 an, während der Mittelwert der Gruppe mit hohem Anfangs-Lie von 5,188 zum Messzeitpunkt t1 auf 4,469 zum Messzeitpunkt t5 abfällt. Der Verlauf ist gegenläufig und statistisch ‚hochsignifikant‘. Man kann in diesem Ergebnis die Tendenz erkennen, dass sich die ‚stärker widerstandsbehaftete‘ Subgruppe im Therapieverlauf mehr öffnet und immer mehr auf Antworten verzichtet, die sie in einem günstigen Licht erscheinen lässt. Hingegen scheint die anfänglich ‚geringer widerstandsbehaftete‘ Subgruppe im Laufe der Therapie immer stärker dahin zu tendieren, sich in einem günstigen Licht darstellen zu wollen, was die Lie-Werte ansteigen lässt.

Im VEV unterscheiden sich zum Messzeitpunkt t3 die Mittelwerte zwischen der Subgruppe „geringeres Anfangs-Lie“ und „höheres Anfangs-Lie“ nur sehr gering (sechs Mittelwertpunkte) und erreicht damit kein Signifikanzniveau. Die beiden Subgruppen schätzen die durch die Therapie erzielte Veränderung im Therapiezeitraum von t1 bis t3 in etwa gleich groß ein. Zum Messzeitpunkt t5, also in der Bewertung der eingetretenen Veränderungen zwischen den Messzeitpunkten t3 und t5, steigt der Mittelwert der Subgruppe „geringeres Lie“ nur unwesentlich um 0,5 Punkte, der Mittelwert der Subgruppe „höheres Lie“ dagegen um zwölf Punkte an. Das heißt, dass die Subgruppe mit dem höheren Anfangs-Lie die Veränderungen im Beobachtungszeitraum etwas größer einschätzen als die Subgruppe mit dem geringeren Anfangs-Lie. Keiner dieser Werte erreicht jedoch ein Signifikanzniveau: Weder unterscheiden sich die Subgruppen zum Zeitpunkt der Messung signifikant, noch zeigen sie, was ihre Therapieeinschätzung anbelangt, im Zeitverlauf eine signifikant unterschiedliche Entwicklung, - noch zeigt sich ein Gesamt-Psychotherapieeffekt dahingehend, dass sich die „Veränderungseinschätzung der Gesamtgruppe“ zwischen den Messzeitpunkten im Ganzen signifikant verändert.

Als Ergebnis bleibt festzuhalten, dass die Ho-Hypothese bestätigt werden konnte, wonach sich nämlich die Gruppen mit „geringem Anfangs-Lie“ und „hohem Anfangs-Lie“ hinsichtlich ihrer Empfänglichkeit für psychotherapeutische Wirkungen nicht signifikant unterscheiden. Insofern man die Stärke des Anfangs-Lie als ein Maß für den Therapiewiderstand ansehen kann, lässt sich das Ergebnis auch so formulieren, dass zwischen der Stärke des Therapiewiderstandes und dem Therapieerfolg im statistischen Sinne kein Zusammenhang nachgewiesen werden konnte. Dennoch ist in diesem Zusammenhang ein deskriptiver und direkter Mittelwertvergleich der klinischen Skalen der Messzeitpunkte t3 und t5 interessant: Während die Mittelwerte der Subgruppe „geringes Anfangs-Lie“ eher stagnieren oder nur gering rückläufig sind, fallen die Mittelwerte der Subgruppe „hohes Anfangs-Lie“ zumindest in den fünf GBB-Skalen und vier klinischen Skalen des MMPI(K) (HY, D, CS und PT) durchweg doch so kräftig, dass sie sich zum Messzeitpunkt t5 in etwa den Werten der Subgruppe „geringeres Anfangs-Lie“ angenähert haben. Interpretativ kann hier trotz nicht nachgewiesener Signifikanz doch von einem weiteren Therapiebedarf und von einem zu erwartenden weiteren Absinken der Mittelwerte ausgegangen werden, wie sich dies zum Messzeit-

punkt t5 auch in der Skala Depression zeigt, in welcher die Gruppe mit dem höheren Anfangs-Lie den Mittelwert der Gruppe mit dem geringeren Anfangs-Lie zum Messzeitpunkt t5 bereits um mehr als einen Mittelwertpunkt unterschritten hat.

#### **10.4.10 Varianzanalyse mit Variablen „Anfänglicher Beschwerdedruck“ (vgl. Tabellen 73 – 75 im Anhang)**

Unsere theoriegeleitete Anfangsvermutung lautete, dass Personen mit ‚starkem Beschwerdedruck‘ durch Psychotherapie erfolgreicher zu behandeln sind als solche, bei denen sich ein eher ‚geringerer Beschwerdedruck‘ findet. Es handelt sich hierbei letztlich um die Überprüfung einer zentralen Grundthese für die prognostische Einschätzung jeder Psychotherapie, dass nämlich der Erfolg von Psychotherapie u.a. auch wesentlich vom Ausmaß des Leidensdrucks abhängt. Wir möchten feststellen, ob sich zwischen der Gruppe mit „hohem anfänglichen Leidens-/Beschwerdedruck“ und der Gruppe mit „niedrigerem anfänglichen Leidens-/Beschwerdedruck“ zu den Messzeitpunkten t1, t3 und t5 ein Unterschied hinsichtlich der Ausprägung in den klinischen Skalen des GBB, MMPI(K) und des VEV zeigt bzw. ob sich diese beiden Leidensdruckgruppen im Zeitverlauf der Psychotherapie hinsichtlich der Ausprägung auf den klinischen Skalen signifikant unterschiedlich entwickeln oder nicht. Wir vermuteten, dass der Psychotherapieerfolg, gemessen in der Veränderung der Ausprägung auf den klinischen Skalen der genannten Messverfahren, um so größer sein wird, je höher der Beschwerdedruck zu Beginn der Therapie ist.

Unsere falsifizierbare Ho-Forschungshypothese lautet, dass es keinen statistisch signifikanten Zusammenhang zwischen dem Ausmaß des Therapieerfolgs und der Höhe des anfänglichen Leidens-/Beschwerdedrucks gibt.

Die Bildung der statistischen Untersuchungsgruppen mit „geringem anfänglichen Beschwerdedruck“ und „hohem anfänglichen Beschwerdedruck“ wurde empirisch, objektiv und wertunabhängig, also ohne Berücksichtigung subjektiver Therapeuteinschätzungen, vorgenommen. Dabei bedienten wir uns der Einschätzungen unserer Patienten selbst, indem wir nämlich ihre Daten in der Beschwerdedruck-Skala des GBB nutzten, um die Gesamtpopulation über den Median in zwei statistische Gruppen aufzuteilen, nämlich in die Gruppe der Patienten, die vom Median der Rangreihe aus gesehen entweder einen ‚höheren‘ oder einen ‚niedrigeren‘ Messwert in der Beschwerdedruckskala aufwiesen.

Die Auswertung erfolgte mittels der Varianzanalyse. Die Betrachtung des Therapieerfolgs der Gesamtgruppe zeigte, dass die Mittelwerte der zusammengefassten Subgruppen zwischen den Messzeitpunkten t1 und t3 in allen fünf GBB-Skalen, in den vier zentralen klinischen Skalen des MMPI(K) (neurotische Trias, bestehend aus Hypochondriasis, Depression, Hysterie sowie Psychasthenie) und in zwei weiteren MMPI(K)-Skalen eine ‚hochsignifikante‘ Senkung erfuhren, was einer Symptomminderung im Sinne einer Gesundung entspricht. In zwei weiteren MMPI(K)-Skalen lässt sich darüber hinaus eine ‚sehr signifikante‘ Mittelwertsenkung nachweisen, eine Skala reagiert noch ‚signifikant‘ und zwei weitere noch ‚tendenziell signifikant‘. Nur die Lie-Skala reagiert ‚nicht signifikant‘.

Der Vergleich der Messzeitpunkte t3 und t5 bescherte uns nicht mehr so starke Mittelwertverbesserungen wie zwischen den Messzeitpunkten t1 und t3. Eine ‚hochsignifikante‘ Mittelwertverbesserung lässt sich nur noch in der Skala Psychasthenie des MMPI(K) nachweisen. Ein ‚sehr signifikant‘ wird lediglich in der

Skala „Herzschmerzen“ des GBB erreicht, ‚signifikant‘ noch beim Beschwerdedruck (GBB), Hypochondriasis, Hysterie und Schizophrenie (MMPI(K)), ‚tendenziell signifikant‘ noch bei drei weiteren MMPI(K)-Skalen. In immerhin drei Skalen des GBB und sechs weiteren Skalen des MMPI(K) lässt sich überhaupt keine signifikante Veränderung mehr nachweisen, auch wenn die Mittelwerte in sieben von neun Skalen fallende und nur in zwei Skalen steigende Tendenz aufweisen. Dies spricht dafür, dass sich die im Sinne einer Gesundung erzielten Veränderungen tendenziell eher zu Beginn als gegen Ende der Therapie ereigneten.

Der Vergleich von t1 zu t5, also gewissermaßen die Überlagerung der gemessenen Veränderungen zwischen den Zeitpunkten t1 und t3 sowie von t3 zu t5 (unter ausschließlicher Berücksichtigung der zu t5 noch beteiligten „N“ Patienten) zeigt, dass in insgesamt 12 Skalen ‚hochsignifikante‘, in drei Skalen ‚sehr signifikante‘, in einer Skala ‚signifikante‘ und in einer weiteren Skala noch ‚tendenziell signifikante‘ Verbesserungen der Symptomwerte nachweisbar sind. Nur die Mittelwerte der Lie-Skala bleiben unterhalb jedweder Signifikanzschwelle.

Betrachtet man den Verlauf der beiden Subgruppen „niedriger Beschwerdedruck“ und „hoher Beschwerdedruck“ im Vergleich, dann zeigt sich, dass sich beide Gruppen zum Messzeitpunkt t1 in 5 Skalen des GBB und in 3 Skalen des MMPI(K) voneinander ‚hoch signifikant‘, in fünf Skalen des MMPI(K) ‚sehr signifikant‘, in keiner Skala ‚signifikant‘, in einer Skala noch ‚tendenziell signifikant‘ und in vier Skalen (Lie, Maskulinität/Femininität, Hypomanie, Soziale Introversion) nicht voneinander unterscheiden, wobei diese ‚Unterschiede‘ durchweg auf eine höhere Symptombelastung der Subgruppe „höherer Beschwerdedruck“ im Vergleich zu der Subgruppe „niedriger Beschwerdedruck“ zurückzuführen sind. Zum Messzeitpunkt t3 ebnen sich diese Unterschiede eher ein. Zu diesem Zeitpunkt ist ein ‚hochsignifikanter‘ Unterschied zwischen den Subgruppen nur noch in den Skalen Erschöpfungsneigung (GBB), Gliederschmerzen (GBB), Beschwerdedruck (GBB) und Hypochondriasis (MMPI/K) zu beobachten. In drei weiteren Skalen ist der Unterschied noch ‚sehr signifikant‘ und in abermals vier weiteren Skalen noch ‚signifikant‘. Dagegen haben sich die beiden Subgruppen in sieben Skalen soweit angeglichen, dass bei diesen kein signifikanter Unterschied mehr messbar ist.

Betrachtet man die Unterschiede der beiden Subgruppen im Zeitverlauf, dann zeigt sich beim GBB ein ‚hoch signifikanter‘ Unterschied der beiden Subgruppen im Zeitraum von t1 zu t3: in der Gruppe ‚geringer Beschwerdedruck‘ fallen die Absenkungen der die Krankheitsbelastung der Patienten darstellenden Mittelwerte deutlich geringer aus als in der Gruppe ‚hoher Beschwerdedruck‘. In den Skalen Erschöpfungsneigung, Magenbeschwerden, Gliederschmerzen und Beschwerdedruck des GBB fallen die Symptomabsenkungen ‚hoch signifikant‘ höher aus als in der Subgruppe „geringer Beschwerdedruck“, - in der Skala Herzschmerzen des GBB erreicht dieser Unterschied noch das Niveau einer einfachen Signifikanz. Wir werten dieses Ergebnis als Hinweis auf einen größeren Therapieerfolg in der Gruppe mit höherem Beschwerdedruck im Gegensatz zu einem geringeren Therapieerfolg in der Gruppe mit geringerem Beschwerdedruck. Insofern deckt sich dieses Ergebnis mit unseren Erwartungen. Allerdings wiederholt sich dieses für den GBB geltende Ergebnis des signifikant unterschiedlichen Verlaufs der untersuchten Subgruppen zwischen den Messzeitpunkten t1 und t3 überraschenderweise nicht in den Skalen des MMPI(K). Hier kann zwar ein therapiebewirkter Abfall der Mittelwerte festgestellt werden, jedoch verhalten sich die Mittelwerte der Subgruppen eher synchron und nehmen größenordnungsmäßig in etwa in gleichartigem Umfang ab.



Bei der Betrachtung des Verlaufs der beiden Subgruppen zwischen den Messzeitpunkten t3 und t5, zeigt sich, dass sich die im Messzeitpunktvergleich von t1 und t3 noch vorhandenen Abbildungsunterschiede zwischen den Messinstrumenten GBB und MMPI(K) verlieren. Beide untersuchten Subgruppen weisen jetzt in allen GBB- und MMPI(K)-Skalen in etwa die gleiche Entwicklung im Sinne von Mittelwertabsenkungen auf. Nur in der Skala Herzschmerzen (GBB) und Schizophrenie (MMPI/K) zeigt sich noch ein signifikanter Unterschied zugunsten eines massiv stärkeren Mittelwertabfalls der Symptombelastung in der Gruppe mit dem „anfänglich höheren Beschwerdedruck“.

Der Vergleich der Messzeitpunkte t1 zu t5 führt zu einer überlagerten Betrachtung der Unterschiede der beiden Subgruppen im Zeitverlauf von t1 zu t3 und von t3 zu t5 mit dem Ergebnis, dass sich wiederum in den GBB-Skalen sehr signifikante Unterschiede zwischen den Verläufen der beiden Subgruppen abbilden, die sich in den MMPI(K)-Skalen so jedoch nicht wiederfinden lassen.

Der Vergleich der Messwerte im VEV zwischen den Messzeitpunkten t3 und t5 zeigt, dass die Unterschiede der Mittelwerte zwischen den Subgruppen „geringerer“ und „höherer anfänglicher Beschwerdedruck“ sehr gering sind und kein Signifikanzniveau erreichen. Zum Messzeitpunkt t3 unterscheiden sich die beiden Subgruppen in ihrer Einschätzung des Therapieerfolgs fast gar nicht (drei Mittelwertpunkte) voneinander, beide Subgruppen schätzen die durch die Therapie erzielte Veränderung im Therapiezeitraum t1 bis t3 in etwa gleich groß ein. Zum Messzeitpunkt t5, also zum Zeitpunkt der Bewertung der eingetretenen Veränderungen zwischen t3 und t5, steigt der Mittelwert der Subgruppe „geringerer Beschwerdedruck“ nur unwesentlich um zwei Punkte, der Mittelwert der Subgruppe „höherer Beschwerdedruck“ nur unwesentlich um acht Punkte an. Das heißt, dass die Subgruppe mit dem höheren Beschwerdedruck die Veränderungen im Beobachtungszeitraum etwas höher einschätzen als die Subgruppe mit dem geringeren Beschwerdedruck. Keiner dieser Werte erreicht jedoch ein noch so geringes Signifikanzniveau: Weder unterscheiden sich die Subgruppen zum Zeitpunkt der Messung signifikant, noch zeigen sie im Zeitverlauf eine signifikant unterschiedliche Entwicklung, - noch zeigt sich dahingehend ein Gesamteffekt, dass sich die „Veränderungseinschätzung der Gesamtgruppe“ zwischen den Messzeitpunkten im Ganzen signifikant verändert.

Zunächst bleibt festzuhalten, dass bei beiden in der Höhe des anfänglichen Beschwerdedrucks sich unterscheidenden Subgruppen signifikante Therapieerfolge gemessen werden konnten. Die Ho-Hypothese, dass es hierbei keinen Zusammenhang zwischen der Höhe des Therapieerfolgs und der Höhe des Leidens-/Beschwerdedrucks gibt, konnte nur teilweise zurückgewiesen werden. Der Unterschied in der „Therapieempfindlichkeit“ zwischen den Subgruppen konnte lediglich in den Skalen des GBB und auch lediglich für den Zeitraum zwischen den Messzeitpunkten t1 und t3 nachgewiesen werden. Für diesen umschriebenen Zeitraum gilt, dass die Gruppe mit dem höheren Beschwerdedruck auch die besseren Therapieergebnisse im Sinne einer empirisch messbaren Symptomminderung aufzuweisen hat. Der Vergleich der Messzeitpunkte t3 und t5 mit Hilfe des GBB zeigte eher einen parallelen Verlauf der beiden Subgruppen. Mit Hilfe des MMPI(K) konnten (bis auf einen einzigen und daher zu vernachlässigenden Messwert) keine signifikant unterschiedlichen Verläufe der Subgruppen festgestellt werden.

Betrachtet man den Gesamtzeitraum zwischen t1 und t5, dann zeigt sich, dass sich die Mittelwerte der beiden Subgruppen in den meisten klinischen Skalen auch noch zum Zeitpunkt t5 deutlich voneinander unterscheiden. Die beiden Subgruppen mit geringem und hohem anfänglichen Leidensdruck haben sich trotz starker Therapieerfolge im Sinne hochsignifikanter Minderungen des Beschwerdedrucks auch nach

drei Jahren Psychotherapie immer noch nicht so stark einander angenähert, dass man sagen könnte, die Subgruppen seien gleich. Auch zum Zeitpunkt t5 haben wir es noch mit massiven Symptomunterschieden zu Lasten der Subgruppe „höherer anfänglicher Beschwerdedruck“ zu tun, was als ein Hinweis darauf zu werten ist, dass zumindest diese Gruppe noch Therapiepotential bzw Therapiebedarf hat. Die Mittelwertabsenkungen dieser Gruppe im Zeitraum zwischen t3 und t5, die zumindest in den fünf GBB-Skalen und in eigentlich allen MMPI(K)-Skalen bis auf die klinisch weniger relevanten Lie, V, C und M/F doch sehr gravierend sind, sprechen dafür, dass diese Entwicklung zum Zeitpunkt t5 noch nicht ihren Sattelpunkt erreicht hat, - eine Fortsetzung der Therapie in diesen Fällen also auch zu diesem Zeitpunkt noch durchaus sinnvoll und notwendig erscheint.

#### **10.4.11 Varianzanalyse mit der Variablen „Anfangsdepression“ (vgl. Tabellen 76 – 78 im Anhang)**

Unsere theoriegeleitete Anfangsvermutung lautete, dass unter Depression leidende Patienten durch Psychotherapie insofern erfolgreich behandelbar sind, als dass die Depressionswerte selbst, aber auch alle anderen klinisch relevanten psychischen Symptomwerte während der Psychotherapie deutlich gemindert werden. Die Verminderung der Symptomwerte messen wir in den klinischen Skalen des GBB, des MMPI(K) und des VEV. Die signifikante Verminderung der Symptomwerte auf diesen Skalen definieren wir als Therapieerfolg. Wir vermuten weiter, dass der Therapieerfolg um so größer sein wird, je höher die Depressionswerte zu Beginn der Therapie sind.

Hieraus lassen sich die negativ umformulierten und falsifizierbaren Ho-Forschungshypothesen ableiten, dass Psychotherapie keine signifikante Veränderung der Messergebnisse in den klinischen Skalen depressiver Patienten bewirkt und dass diesbezüglich auch kein signifikanter Unterschied zwischen der Gruppe der anfänglich hochdepressiven und der anfänglich geringer depressiven Subgruppe festzustellen ist.

Die Bildung der statistischen Gruppen erfolgte wiederum empirisch, objektiv und ohne Beteiligung durch die Therapeuten. Sie wurde letztlich von den Patienten dadurch selbst vorgenommen, dass sie sich zum Messzeitpunkt t1 durch ihr Antwortverhalten auf der Depressivitätsskala des MMPI(K) in eine Depressivitätsrangreihe brachten, die über den Median die Möglichkeit zur Bildung zweier statistischer Gruppen, nämlich derjenigen mit „niedrigerer“ und „höherer Anfangsdepression“ schuf. Dieses Verfahren entspricht empirischen Standards, enthebt uns jedweder Diskussion über intersubjektiv nicht überprüfbare Therapeuteinschätzungen, macht uns jedoch abhängig von der Validität des MMPI(K) in der Depressivitätsskala.

Die Varianzanalyse erbrachte folgende Ergebnisse: Die Mittelwerte der zusammengefassten Subgruppen erfahren zwischen den Messzeitpunkten t1 und t3 in allen klinischen Skalen des GBB und in sieben klinischen Skalen des MMPI(K) eine hochsignifikante Veränderung im Sinne einer Werteminderung, was einer Symptomverbesserung gleichkommt. In zwei Skalen des MMPI(K) lässt sich eine sehr signifikante Mittelwertsenkung, in zwei weiteren Skalen eine signifikante und in einer weiteren Skala eine immerhin noch tendenziell signifikante Mittelwertsenkung feststellen. Nur die Lie-Skala des MMPI(K) reagiert nicht signifikant, was heißt, dass in dieser Skala keine Veränderung durch Psychotherapie stattgefunden hat.

Betrachtet man die Mittelwertveränderungen zwischen den Messzeitpunkten t3 und t5, dann lassen sich zwar noch ‚signifikante‘ Mittelwertveränderungen im Sinne von Symptomverbesserungen in den diversen klinischen Skalen nachweisen, jedoch erreichen diese Veränderungen nicht mehr das große Ausmaß wie zwischen den Messzeitpunkten t1 und t3. Nur noch in der Skala Psychasthenie wird eine ‚hochsignifikante‘ Veränderung, in zwei weiteren klinischen Skalen eine ‚sehr signifikante‘ Veränderung, in sechs Skalen eine noch ‚signifikante‘, in neun Skalen jedoch keinerlei signifikante Veränderung mehr gemessen. Ein auf der Ebene der Beschreibung vorgenommener Vergleich der Mittelwerte zu den Messzeitpunkten t3 und t5 zeigt, dass in nur zwei Skalen ein Mittelwertanstieg, in allen anderen Skalen jedoch ein deutlicher Mittelwertabfall feststellbar ist, auch wenn dieser Abfall nicht mehr immer als ‚signifikant‘ zu bezeichnen ist.

Betrachtet man den Veränderungszeitraum zwischen t1 und t5, so ist nach dem vorher Gesagten zu erwarten, dass die Zusammenführung der Mittelwertveränderungen im Sinne einer Überlagerung der Messergebnisse von t1 zu t3 und von t3 zu t5 zu einer noch deutlicheren Abbildung des Therapieerfolgs im Sinne von empirisch messbarer Symptomminderung führt.

Entsprechend überrascht es nicht, dass in insgesamt 13 klinischen Skalen ein ‚hochsignifikanter‘ positiver Therapieerfolg, in zwei weiteren Skalen ein ‚sehr signifikanter‘ Erfolg, in einer Skala ein ‚signifikanter‘ und in einer weiteren Skala ein noch ‚tendenziell signifikanter‘ Erfolg nachweisbar ist. Nur die Lie-Skala reagiert nicht signifikant.

Betrachtet man die Unterschiede zwischen den Subgruppen „geringere Anfangsdepression“ und „höhere Anfangsdepression“ zum Zeitpunkt t1, dann zeigt sich, dass sie sich zu diesem Messzeitpunkt in vier Skalen des GBB und elf Skalen des MMPI(K), also in insgesamt 15 Skalen ‚hochsignifikant‘, in einer Skala des GBB ‚sehr signifikant‘ und in zwei Skalen des MMPI(K) ‚nicht signifikant‘ voneinander unterscheiden. Diese deutliche Unterscheidung der Subgruppen geht zum Messzeitpunkt t3 weitgehend verloren. Dies hat mit der vor allem in der Subgruppe der „höheren Anfangsdepression“ stattgefundenen Absenkung der Symptombelastung, ausgedrückt über das Abfallen der erzielten Skalenmittelwerte, zu tun. So unterscheiden sich die beiden Subgruppen zum Zeitpunkt t3 in keiner Skala mehr ‚hochsignifikant‘, in nur sieben Skalen noch ‚sehr signifikant‘, in noch drei Skalen ‚signifikant‘ und in drei Skalen noch ‚tendenziell signifikant‘ voneinander. In fünf Skalen zeigen die Mittelwerte der beiden Subgruppen keine signifikante Abweichung mehr voneinander.

Die Betrachtung des Verlaufs der beiden Subgruppenunterschiede in der Zeitabfolge zwischen den Messzeitpunkten t1 und t3 zeigt eine große Verlaufsparallelität: beide Subgruppen zeigen in den klinischen Skalen Mittelwertveränderungen in gleicher Richtung, nämlich in Richtung „fallender Mittelwerte“, auch wenn das Ausmaß dieser Mittelwertabfälle unterschiedlich ausfällt. In beiden Subgruppen lässt sich somit ein Therapieerfolg nachweisen, nur dass dessen Ausmaß in beiden Subgruppen unterschiedlich ist. In sechs Skalen des GBB und einer Skala des MMPI(K) zeigt sich eine signifikante Abweichung in dem Sinne, dass die Subgruppe mit der höheren Anfangsdepression auch einen signifikant höheren Mittelwertabfall im Sinne einer Symptomminderung (Therapieerfolg) hat, als dies in der Gruppe mit der niedrigeren Anfangsdepression festzustellen ist. In zwei Skalen des MMPI(K) zeigt sich eine noch ‚tendenzielle‘ Signifikanz, ebenfalls im Sinne eines höheren Mittelwertabfalls der Subgruppe mit der höheren Anfangsdepression. In den restlichen zwei Skalen des GBB und den sieben MMPI(K)-Skalen zeigt sich hingegen keinerlei signifikante unterschiedliche Entwicklung, was heißt, dass zwar in beiden Gruppen ein großer

Therapieerfolg nachweisbar ist, dass dieser Therapieerfolg sich jedoch in den beiden Subgruppen nicht signifikant unterschiedlich auswirkt.

Im weiteren Vergleich der Messzeitpunkte t3 zu t5 zeigt sich, dass die Subgruppe mit der geringeren Anfangsdepression in mehreren Skalen eine deutlich geringere Mittelwertabsenkung erreicht als die Subgruppe mit der höheren Anfangsdepression. In der Skala Psychasthenie zeigt sich diese Tendenz hochsignifikant: Während der Mittelwert der Symptombelastung in der Subgruppe „geringe Anfangsdepression“ nur noch von 9,625 auf 9,500 abfällt, zeigt sich bei der Gruppe der „höheren Anfangsdepression“ ein Abfall von 14,839 auf 11,032 Punkte. Bei einigen Skalen zeigt sich jedoch auch eine gegenläufige Entwicklung beider Subgruppen, und zwar derart, dass die Subgruppe „geringe Anfangsdepression“ nach der deutlichen Symptomminderung von t1 zu t3 jetzt von Messzeitpunkt t3 zu Messzeitpunkt t5 mit einer Symptomzunahme reagiert, während die Subgruppe „hohe Anfangsdepression“ weiterhin eine Symptomabnahme zu verzeichnen hat. Dieser signifikant gegenläufiger Verlauf beider Subgruppen zeigt sich in der GBB-Skala Gliederschmerzen und den MMPI(K)-Skalen Depression, Validity, Psychopathie, Schizophrenie, und Soziale Introversion.

In einigen Fällen kommt es aber auch bei der Subgruppe der „höheren Anfangsdepression“ nicht zu einer weiteren Abnahme, sondern zu einer eher leichten Zunahme der Mittelwerte, nämlich in den MMPI(K)-Skalen Lie, Correction und Masculinity/Femininity, - wobei es sich bei allen drei Skalen jedoch nicht um ausgesprochen klinische Skalen handelt.

Bezogen auf den Subgruppenvergleich der beiden Messzeitpunkte t1 und t5 kommt es zu einer Überlagerung der jeweiligen Verlaufseffekte von t1 zu t3 und von t3 zu t5. Von daher ist es nicht überraschend zu sehen, dass die Therapie in der Subgruppe der „höher Anfangsdepressiven“ weit stärkere positive und in Richtung Gesundung weisende Effekte hatte als dies in der Subgruppe der „geringer Anfangsdepressiven“ zu verzeichnen war: In immerhin drei Skalen des MMPI(K) (Depression, Psychasthenie und Schizophrenie) lässt sich diesbezüglich ein ‚hochsignifikanter‘ Unterschied zwischen den beiden Subgruppen nachweisen. In den Skalen Beschwerdedruck (GBB) und den MMPI(K)-Skalen Hypochondriasis, Hysterie, Validity, Psychopathie und Paranoia ist der Verlaufsunterschied zwischen den beiden Subgruppen noch ‚sehr signifikant‘, in zwei weiteren Skalen ‚signifikant‘ und in drei weiteren Skalen noch ‚tendenziell signifikant‘ nachweisbar. Keinen signifikanten Verlaufsunterschied der Subgruppen fanden wir lediglich in den Skalen Gliederschmerzen beim GBB sowie den MMPI(K)-Skalen Lie und Hypomanie.

Im VEV-Vergleich von t3 zu t5 sind die Unterschiede der Mittelwerte zwischen der Subgruppe „geringere Anfangsdepression“ und „höhere Anfangsdepression“ sehr gering und erreichen kein Signifikanzniveau. Die beiden Subgruppen unterscheiden sich fast gar nicht (vier Mittelwertpunkte) voneinander, beide Subgruppen schätzen die durch die Therapie erzielte Veränderung im Therapiezeitraum von t1 bis t3 in etwa gleich groß ein. Zum Messzeitpunkt t5, also in der Bewertung der eingetretenen Veränderungen zwischen t3 und t5, steigen der Mittelwert der Subgruppe „geringere Anfangsdepression“ nur unwesentlich um einen Punkt, der Mittelwert der Subgruppe „höhere Anfangsdepression“ dagegen um 14 Punkte an. Das heißt, dass die Subgruppe mit der höheren Anfangsdepression die Veränderungen im Beobachtungszeitraum etwas höher einschätzt als die Subgruppe mit der geringeren Anfangsdepression. Keiner dieser Werte erreicht jedoch ein Signifikanzniveau: Weder unterscheiden sich die Subgruppen zum Zeitpunkt der Messung signifikant, noch zeigen sie im Zeitverlauf eine signifikant unterschiedliche Entwick-

lung, - noch zeigt sich ein Gesamt-Psychotherapieeffekt dahingehend, dass sich die „Veränderungseinschätzung der Gesamtgruppe“ zwischen den Messzeitpunkten im Ganzen signifikant verändert.

Die Ho-Hypothese, dass Psychotherapie an depressiven Patienten keine signifikante Auswirkung auf die Messergebnisse der klinischen Skalen der Patienten hat, konnte falsifiziert/zurückgewiesen werden. Es konnten statt dessen hochsignifikante Veränderungen im Sinne von Symptomverbesserungen in den meisten klinischen Skalen des GBB und des MMPI(K) nachgewiesen werden. Auch die Hypothese, dass es hierbei keinen Unterschied zwischen der Subgruppe mit der geringeren und der Subgruppe der höheren Anfangsdepression gäbe, konnte erfolgreich falsifiziert/zurückgewiesen werden. Es zeigte sich, dass Langzeitpsychotherapie bei der Subgruppe der höher depressiven Patienten auch dann noch signifikante Zustandsverbesserungen bewirkt, wenn diese Entwicklung sich in der Gruppe der geringer depressiven Patienten bereits verlangsamt oder leicht umgekehrt hat, - der Sattelpunkt also durchschritten ist.

## **10.5 Zusammenfassung und Kommentierung der Effektivitätsstudie**

Die von Anfang bis Ende des Untersuchungszeitraumes um die Hälfte reduzierte Teilnehmerzahl scheint bei Untersuchungen wie der unsrigen nicht ungewöhnlich zu sein. Dies kann damit zusammenhängen, dass entweder der gewünschte Behandlungserfolg sich schon frühzeitig einstellte, dass die Behandlung abgebrochen wurde oder die Patienten sich gar nicht mehr oder nicht mehr anweisungsgerecht an der Untersuchung beteiligt haben.

Unsere Patienten sind hauptsächlich 30 bis 40 Jahre alt, Frauen und Männer sind in einem Verhältnis von ungefähr 70:30 vertreten. Annähernd repräsentativ sind unsere Zahlen für die Altersgruppen 21–30 und 41–50. Über die Hälfte der Stichprobe ist Beamter oder Angestellter von Beruf. Schulabschluss ist bei der Hälfte der Stichprobe das Abitur.

Diagnostisch handelt es sich vorrangig um Patienten mit der Diagnose „depressive Neurose“. Die Diagnose „narzißtische Persönlichkeitsstörung“ ist seltener vertreten, andere Diagnosen kommen nur vereinzelt vor. Im Klassifikationsschema des ICD-10 ausgedrückt ergibt sich ein Schwergewicht bei „affektive Störung“ (depressive Neurose), „neurotische Störung“ und „spezifische Persönlichkeitsstörung“. Sie machen je ungefähr 1/4 der von den Therapeuten gestellten Ursprungsdiagnosen aus.

Patienten mit schwereren (frühen) Störungen bleiben länger in Behandlung, Patienten mit neurotischen Störungen beenden ihre Behandlung früher. Über die Gründe dafür geben unsere Daten keine verlässlichen Hinweise. Es gibt aber auch Anzeichen dafür, dass die schwerer gestörten Patienten mit Persönlichkeitsstörungen die Behandlung frühzeitig abgebrochen haben, während die weniger kranken Patienten mit der Diagnose „neurotische“ bzw. „affektive Störung“ schneller nicht mehr behandlungsbedürftig waren und die Behandlung deshalb beendet haben.

Die t-Tests in der unbehandelten Kontrollgruppe zeigen, wie zu erwarten war, durchweg nicht-signifikante Veränderungen zwischen t1 und t2, außer dem signifikanten Unterschied in der MMPI-Skala Hysterie, in der die Werte von t1 nach t2 überraschenderweise auch ohne Behandlung signifikant abnehmen. Auch von t2 nach t3 sind die Veränderungen in den Mittelwerten der GBB- und MMPI(K)-Skalen gering. Im VEV zeigen sich keine Veränderungen im Erleben und Verhalten im letzten Halbjahr der Wartezeit.

Legt man den gesamten Einjahreszeitraum der Kontrollgruppenlaufzeit zugrunde, so ergibt sich zwischen t1 und t3 nun auch für GBB Skala 4 (Herzbeschwerden) und wiederum für MMPI-Hysterie eine nicht erwartete signifikante Abnahme, während in den anderen (vier GBB- und drei MMPI-) Skalen die Unterschiede zwischen t1 und t3 statistisch nicht von Bedeutung sind.

Neben dem t-Test haben wir die über die rechnerischen Möglichkeiten des t-Tests hinausgehende Varianzanalyse (vgl. Kap. 5.13.2) gerechnet, um eine Signifikanzprüfung des unterschiedlichen Verlaufs von Untersuchungs- (UG) und Kontrollgruppe (KG) in einem einzigen zusammenfassenden Rechengang durchführen zu können. Dabei konnte in den meisten klinischen Skalen nachweisen werden, dass in der UG signifikant höhere Mittelwertabsenkungen auftreten als in der KG: in den drei GBB-Skalen Erschöpfungsneigung, Magenbeschwerden und Beschwerdedruck zeigten sich „signifikante“ bzw. „sehr signifikant“ größere Symptomminderungen in der UG als in der KG. Ähnlich eindeutig ist das Ergebnis auch in den vier als „Neuroseindikator“ besonders relevant gewerteten Skalen des MMPI(K), von denen die Skalen Depression und Psychasthenie einen „sehr signifikanten“, die Skala Hypochondrie einen „signifikanten“ und die Skala Hysterie immerhin noch einen „tendenziell signifikanten“ Unterschied zwischen der UG und der KG zugunsten der UG anzeigen. „Signifikant“ zeigen darüber hinaus auch noch die neurosetheoretisch weniger bedeutsamen MMPI(K)-Skalen Lie und Psychopathic Deviate sowie sogar „sehr signifikant“ die Skala Schizoidie an. Dieses Ergebnis ist insgesamt gesehen als Nachweis für die Wirkung von Psychotherapie in unserer UG sehr überzeugend, auch wenn die beiden GBB-Skalen Gliederschmerzen und Herzschmerzen wie auch die MMPI(K)-Skalen Paranoia, Hypomanie und Soziale Introversion „nicht signifikant“ anzeigen.

Die t-Test-Betrachtung des Kontrollgruppenverlaufs macht dennoch deutlich, dass sich die Annahme nicht uneingeschränkt aufrechterhalten lässt, in der Wartezeit zeigten sich keine Verbesserungen, so wie sie durch Psychotherapie erzielt werden können. Verbesserungen treten auch ohne Psychotherapie – als Spontanremissionen bekannt – allerdings in bescheidenerem Umfang, begrenzt auf ein bestimmtes Störungsbild (Hysterie und Herzbeschwerden) und von unbekannter Stabilität auf.

Die t-Tests in der Therapiegruppe ergeben für das erste Halbjahr der psychotherapeutischen Behandlung eine hochsignifikante Abnahme aller Mittelwerte in den fünf GBB-Skalen und den Skalen der neurotischen Trias (Hypochondrie, Hysterie, Depression) sowie Psychasthenie des MMPI(K). Es sind demnach sowohl die Körpersymptome als auch die neurotischen Störungen nach dem ersten halben Jahr psychotherapeutischer Behandlung substantiell gebessert.

Im zweiten Halbjahr der Behandlung (von t2 bis t3) schreitet die Gesundung der Patienten weiter fort, jedoch nicht mehr wie im ersten Halbjahr durchgehend mit statistisch hoch signifikanten Mittelwertunterschieden. Unter den Körpersymptomen geht es mit den Magenbeschwerden am schwierigsten voran im zweiten Behandlungshalbjahr, während Erschöpfungsgrad und Beschwerdedruck weiterhin deutlich abnehmen. Das heißt, die Patienten verspüren weniger durch körperliche Beschwerden bedingten Leidensdruck und kommen wieder vermehrt zu Kräften, was sowohl der starke Rückgang von Depression (MMPI(K)-Depression) als auch das im VEV gemessene veränderte Erleben und Verhalten zeigt.

Im zweiten Behandlungsjahr reduziert sich die Wirksamkeit der psychotherapeutischen Behandlung in den meisten untersuchten Dimensionen. In MMPI(K)-Hysterie und MMPI(K)-Hypochondrie sind die Mittelwertsunterschiede sehr signifikant und auch noch im VEV. Die Entwicklung in der Hypochondrie-Skala des MMPI(K) macht aber auch deutlich, dass es Störungen gibt, bei denen sich erst nach längerer Zeit

eine Veränderung nachweisen lässt. Im Unterschied dazu gehen die Hysterie-Werte kontinuierlich zurück.

Der Vergleich von Ende des zweiten mit Ende des dritten Behandlungsjahres ergibt noch einen sehr signifikanten Unterschied in MMPI(K)-Psychasthenie, einen signifikanten in GBB-Herzbeschwerden und tendenziell signifikanten noch in MMPI(K)-Depression. Alle anderen Unterschiede, auch die im VEV, sind nicht signifikant. Unsere Ergebnisse stimmen in der Tendenz mit Befunden aus dem Heidelberger Kataneseprojekt überein, dass das Wirkungszenit bei 160 Stunden und 2 1/2 Jahren Behandlungsdauer liegt (Kordy et al., 1989).

Folgende Erklärungen bieten sich für den Wirkungsrückgang nach drei Jahren Behandlung an:

Die positiven Effekte werden bei der Mittelwertberechnung durch einen größer gewordenen Anteil von Patienten, die sich nicht verbessern oder sogar verschlechtern, aufgehoben. Hier wäre eine differenzierte, qualitative Forschung nötig, um Aufschluss über die Gründe zu bekommen.

Nach Howard et al. (1986) ist mit zunehmender Stundenanzahl eine zunehmende Dosis (=Behandlungsstunden) notwendig, um weitere Wirkung zu erzielen. Das durchschnittliche Plus von 61 Stunden im dritten Behandlungsjahr (s.Tab.56) reicht demzufolge nicht aus, um zu diesem Zeitpunkt noch Wirkung messbar (mit unseren Instrumenten) zu erfassen. Denkbar ist, dass „feinere“ Messinstrumente strukturelle Veränderungen, die eine lange Entwicklungszeit brauchen, erfassen könnten.

Eine über alle Messzeitpunkte kontinuierliche Entwicklung in allen Dimensionen lässt sich nicht nachweisen, vielmehr haben die einen ihren Endpunkt schon vor dem Ende des Beobachtungszeitraumes (nach zweijähriger Behandlung die GBB-Symptomcluster Erschöpfungsneigung, Gliederschmerzen, Magenbeschwerden, Beschwerdedruck, MMPI(K)-Hysterie und -Hypochondriasis sowie VEV), andere können sich bis zum Ende des dritten Jahres verbessern, nachdem sie im zweiten Behandlungsjahr stagnierten (MMPI(K)-Psychasthenie und -Depression, GBB-Herzbeschwerden).

Im direkten Vergleich zwischen den Ergebnissen von Kontroll- und Untersuchungsgruppe nach einem halben Jahr schneidet die Untersuchungsgruppe in mehreren Skalen statistisch signifikant besser ab als die Kontrollgruppe, d.h. dass sich auf der Basis des Kontrollgruppenvergleichs in diesen Skalen mit einiger Wahrscheinlichkeit die Verbesserung auf die psychotherapeutische Behandlung zurückführen lässt und nicht gleichsam zufällig entstanden ist. Es sind dies der VEV, GBB-Gliederschmerzen und GBB-Beschwerdedruck, bei denen die Unterschiede (tendenziell) signifikant sind. In den anderen Skalen zeigen sich keine statistisch bedeutsamen Unterschiede im direkten Vergleich zwischen Untersuchungsgruppe und Kontrollgruppe nach einem halben Jahr. Es fällt noch auf, dass die Signifikanzen in den MMPI(K)-Skalen geringer als in den meisten GBB-Skalen ausfallen.

Erweitert man den Vergleichszeitraum zwischen Untersuchungsgruppe und Kontrollgruppe auf ein Jahr, so werden die Unterschiede zwischen beiden Gruppen deutlicher. Alle GBB-Skalen in der Untersuchungsgruppe weisen nun mindestens signifikant verbesserte Mittelwerte auf im Vergleich zur Kontrollgruppe. Auch im VEV ist der Unterschied sehr signifikant. Im MMPI(K) sind signifikante Unterschiede zwischen Untersuchungsgruppe und Kontrollgruppe in den Skalen Hypochondriasis und Schizoidie feststellbar. In der Untersuchungsgruppe sind die Mittelwerte in den anderen Skalen jeweils niedriger als in der Kontrollgruppe, aber nicht statistisch signifikant.

Es zeigt sich auch in diesen Befunden, dass Veränderung von Störungen, die tiefer in der Persönlichkeit verankert sind, eine längere psychotherapeutische Behandlung als die Verbesserung psychosomatischer

Symptome benötigt. Leichte Verbesserungen können hier schon durch wenige psychotherapeutische Kontakte erzielt werden; deutliche Wirksamkeit bei den psychosomatischen Symptomen zeitigt die psychotherapeutische Behandlung im Vergleich zur einer unbehandelten Kontrollgruppe nach einem Jahr Behandlung.

Die Wirksamkeit der Behandlung lässt sich auch im globalen Effektivitätsmaß (Effektstärke) zeigen. Die Effektstärke in der Untersuchungsgruppe erhöht sich bis zum Ende des zweiten Jahres nach Behandlungsbeginn, um dann bis zum Ende des dritten Jahres etwas zurückzugehen. Der Zuwachs an Effektivität nimmt mit zunehmender Behandlungsdauer ab. Der Effektivitätszuwachs, ausgedrückt in der Effektstärke, in den letzten beiden Jahren der Behandlung entspricht ungefähr dem Zuwachs an Effektivität im zweiten Behandlungshalbjahr. Die Effektstärkenmaße in der Kontrollgruppe für den Zeitraum t1 bis t3 bleiben erwartungsgemäß erheblich unter denen der Untersuchungsgruppe für denselben Zeitraum zurück.

Erfahrene und unerfahrene (Kriterium ist das Abschlussjahr der analytischen Weiterbildung) Therapeuten erzielen bei ihren Behandlungen in der dreijährigen Laufzeit der Untersuchung eine kaum voneinander unterschiedene Effektstärke. Das gleiche gilt für beide an der Untersuchung beteiligten Berufsgruppen, Arzt und Psychologe.

Ein weiteres Kriterium für die Wirksamkeit der von uns untersuchten individualpsychologisch-psychoanalytischen Psychotherapien ist die externe Validierung an den von den Testautoren mitgeteilten Ergebnissen der Vergleichsgruppen.

Der Vergleich der GBB-Ergebnisse nach drei Jahren (t5) mit denen der Eichstichprobe von Brähler u. Scheer (1995): Auf Skalenniveau ist der Abstand zur Eichstichprobe im Grad der Erschöpfung und des Beschwerdedrucks relativ groß, Magenbeschwerden, Gliederschmerzen und Herzbeschwerden haben dagegen so weit abgenommen, dass sie sich von den Gesunden der Eichstichprobe nur noch unwesentlich unterscheiden. Bei Betrachtung der Entwicklung auf der Basis der Einzelbeschwerden zeigt sich, dass in  $\frac{3}{4}$  aller Einzelbeschwerden des GBB nach dreijähriger Behandlungszeit das Gesundheitsniveau der Normalbevölkerung erreicht bis überschritten werden konnte.

Die Wirksamkeit der von uns untersuchten individualpsychologisch-psychoanalytischen Psychotherapie lässt sich auch als Veränderung zu mehr Gelassenheit, Entspannung und Optimismus beschreiben. Im VEV zeigen sich schon nach einem Jahr Behandlung statistisch signifikante Verbesserungen, die sich bis nach zwei Jahren fortsetzen. Am Ende des dritten Behandlungsjahres kann ein weiterer Zuwachs an Gelassenheit und Optimismus nicht festgestellt werden. Die Mittelwerte nach dem zweiten und dritten Behandlungsjahr zeigen nach den Kriterien dieses Verfahrens (s. Tab. 6) eine statistisch signifikante Verbesserung auf dem 0.1 %-Niveau, nach dem ersten Jahr auf dem 1.0 %-Niveau. Im Vergleich zur Therapiegruppe von Zielke (s. Tab. 35) sind unsere Werte schlechter, jedoch ist die sehr unterschiedliche diagnostische und demografische Zusammensetzung der Gruppen zu berücksichtigen.

Im MMPI(K) zeigt der Vergleich unserer 13 Skalenmittelwerte mit den von Gehring u. Blaser (1983) mitgeteilten, dass unsere Patienten nach dreijähriger Behandlung nicht den Grad an Gesundheit wie die Gesunden-Vergleichsgruppe erreichen. Allerdings sind auch hier Zweifel an der Vergleichbarkeit der Gruppen angebracht (s. Kommentar zu Tab. 25).



Die Patienten neigen dazu, sich zum Zeitpunkt vor der Behandlung etwas kranker zu schildern als sie wirklich sind. Dies machen die Kontrollskalen des MMPI(K) deutlich. Tiefenpsychologisch betrachtet sind unsere Patienten bemüht, als Kranke gesehen und akzeptiert zu werden.

Dass unsere Patienten in der Untersuchungsgruppe zu Beginn der Behandlung tatsächlich krank sind, zeigt sich nicht nur an der Tatsache, dass sie eine psychotherapeutische Behandlung aufsuchen, sondern auch darin, dass sie im Vergleich zu der Gesamtgruppe "Gesunde" bei Gehring und Blaser deutlich höhere Mittelwerte aufweisen, wobei der Unterschied besonders groß bei der Skala Depression ausfällt. Allerdings ist der Störungsgrad unserer Patienten im Vergleich zu den in der Handanweisung (Gehring u. Blaser, 1983) befindlichen Störungsgruppen (Patienten zum großen Teil aus der psychiatrischen Klinik) geringer.

Krankheitsnachweise zu Behandlungsbeginn sind weiterhin Erschöpfungsneigung, Magenbeschwerden und Gliederschmerzen unserer Patientengruppe, die ähnlich hoch wie in der Patientenstichprobe von Brähler u. Scheer (1995) sind und - allerdings weniger deutlich - Herzbeschwerden und Beschwerdedruck.

Zum Behandlungsbeginn sind mehr als 4/5 der Patienten davon überzeugt, dass ihre Beschwerden überwiegend oder ausschließlich seelisch bedingt sind, nach drei Jahren sind es etwas mehr als 2/3. Möglicherweise ist dieser hohe Prozentsatz von seelischer Bedingtheit am Anfang der Behandlung bei stationären Patienten nicht zu finden, jedenfalls legen dies die Ergebnisse von Kammerer (1998) nahe. Die Erfahrung mit der psychotherapeutischen Behandlung und dem, was sie (nicht) bewirken kann, hat einige unserer Patienten zur Überzeugung kommen lassen, die Beschwerden könnten nicht seelisch, sondern nur körperlich bedingt sein.

Die Therapiewirksamkeit in verschiedenen Subgruppen (Ergebnisse der Varianzanalysen):

### *Diagnosegruppen*

Die mit Hilfe des ICD-10 gebildeten drei Diagnosegruppen haben zu Beginn der Behandlung ein ähnlich hohes Ausgangsniveau, es gibt keine signifikanten Unterschiede zwischen den Gruppen. Auch in der Entwicklung bis zum Ende des ersten Behandlungsjahres zeigt sich kein signifikanter Unterschied zwischen den drei Diagnosegruppen. Tendenziell scheinen sich Patienten mit der Diagnose „neurotische“ und „affektive Störung“ im ersten Jahr schneller zu entwickeln als Patienten mit der Diagnose „spezifische Persönlichkeitsstörung“.

Die Entwicklung der Diagnosegruppen nach dem ersten bis zum Ende des Beobachtungszeitraumes nach drei Jahren macht signifikante Unterschiede deutlich bei GBB-Gliederschmerzen und bei MMPI(K)-Psychasthenie. Auffällig, allerdings nicht signifikant, wirkt die Behandlung besser bei Patienten mit „affektiver Störung“ und „spezifischer Persönlichkeitsstörung“ als mit „neurotischer Störung“, was in den vier MMPI(K)-Skalen, in GBB-Beschwerdedruck und auch im VEV zum Ausdruck kommt.

Auch die Veränderung der drei Diagnosegruppen im gesamten Beobachtungszeitraum von Behandlungsbeginn bis zum Ende des dritten Behandlungsjahres macht gewisse Unterschiede in der Entwicklung deutlich. Es profitieren vor allem die Depressiven (affektive Störung, Diagnosestellung zu t1) von der Behandlung (sichtbar in GBB-Beschwerdedruck und MMPI(K)-Psychasthenie).

### *Geschlecht der Patienten*

Frauen klagen zu Beginn der Behandlung mehr über Gliederschmerzen und entwickeln mehr konversionsneurotische Symptome (MMPI[K]-Hysterie) als Männer. Der durch die Behandlung erfolgte Rückgang der Symptome hat für beide Geschlechter keinen auffällig verschiedenen Verlauf, tendenziell vielleicht am ehesten noch in der Depressionsentwicklung, wo die Frauen von einem etwas höheren Niveau ausgehend sich zu mehr relativer Depressionsfreiheit verändern.

Nach einem Jahr individualpsychologisch-psychoanalytischer Psychotherapie weisen Frauen einen deutlich höheren Veränderungswert (VEV) hin zu mehr Entspannung und Optimismus auf als Männer. Bis zum Ende des dritten Behandlungsjahres sind die Unterschiede zwischen Männern und Frauen nur in der Skala Psychasthenie signifikant, Frauen gesunden demnach in der Langzeitbehandlung unter dem Aspekt von Zwangsverhalten und phobischen Gedanken (Skala Psychasthenie) mehr als Männer.

Legt man den gesamten Beobachtungszeitraum von drei Jahren zugrunde und nur die reduzierte Grundgesamtheit der Patienten, von denen zu t1 und t5 Daten vorliegen, so erzielen die Frauen in den Hauptneuroseindikatoren des MMPI(K) eine größere Therapiewirkung als die Männer. Bei den somatoformen Beschwerden scheinen die Frauen, was den Beschwerdedruck angeht, etwas mehr von der Behandlung zu profitieren. Insofern die somatoformen Beschwerden im Vordergrund stehen, lässt sich eine unterschiedliche Entwicklung von Männern und Frauen in der dreijährigen Behandlungszeit nicht deutlich nachweisen.

### *Alter der Patienten*

Gliederschmerzen sind zu Beginn der Behandlung das Problem der über 40Jährigen, sonstige altersbedingte Unterschiede liegen zu diesem Zeitpunkt nicht vor.

Im ersten Behandlungsjahr scheinen die über 40Jährigen in fast allen Bereichen am stärksten von der Behandlung zu profitieren. Im Zwei-Jahreszeitraum (nach dem ersten bis nach dem dritten Jahr) gibt es widersprüchliche Tendenzen und Hinweise auf altersbedingte Unterschiede: Einerseits ist eine positive Wirkung in der jüngeren (21–30 Jahre) und der älteren (über 40) Gruppe in ihrem Erleben und Verhalten (VEV) festzustellen, andererseits zeigen sich in den GBB- und MMPI(K)-Skalen deutlicher Langzeiteffekte nur bei der Gruppe der über 40Jährigen und am wenigsten bei den Jüngeren, bei denen die längere Therapiedauer in beiden Testverfahren keinen gut nachweisbaren zusätzlichen Nutzen bringt.

Auch im Drei-Jahresvergleich (von t1 bis t5) sprechen einige Hinweise dafür, dass die Altersgruppe der über 40Jährigen von der dreijährigen Behandlungszeit am stärksten profitiert hat.

### *Dyaden Therapeut/in – Patient/in*

In der Dyade Therapeut/Patientin befinden sich zum Behandlungsbeginn die relativ krankesten Patienten (mit den höchsten Mittelwerten in GBB und MMPI(K)). Bis zum Ende des ersten Behandlungsjahres zeigt sich nur in der Skala GBB-Magenbeschwerden ein (tendenziell) signifikanter Unterschied zwischen den verschiedenen Dyaden. Hier machen Therapeut/Patientin die größten Fortschritte, während Therapeutin/Patient nur eine geringe Veränderung hervorbringen. Im Zweijahreszeitraum (t3-t5) sprechen Hinweise für die geringe Wirksamkeit von Therapeut/Patient und die Wirksamkeit von Therapeutin/Patientin und Therapeut/Patientin.

Im Drei Jahreszeitraum (auf der Basis der drei Jahre in Behandlung befindlichen Patienten), zeigt sich zum Zeitpunkt t1 in allen Skalen, in dreien signifikant, dass in Dyade Therapeut/Patientin kränkere Patienten sind als in den anderen Dyaden.

Frauen verändern sich, wenn man den gesamten Beobachtungszeitraum einbezieht, in den meisten Skalen, besonders in Depression und Hysterie, mehr als Männer, dabei spielt das Geschlecht des Therapeuten keine Rolle.

### *Sitzungsfrequenz*

Zum Behandlungsbeginn sind die beiden Frequenz-Untergruppen (niederfrequent/höherfrequent) nicht unterschiedlich krank.

Die festgestellte Wirksamkeit der Behandlung ist im ersten Jahr, unabhängig von der Stundenanzahl, nachweisbar: Ob nieder- oder höherfrequent ist für den Veränderungsgrad ohne Bedeutung.

Zwei Jahre nach Behandlungsbeginn wurde die Gesamtgruppe wiederum in zwei gleich große Gruppen über dem Median geteilt: Der Vergleich der Gruppe mit eher weniger Stunden (niederfrequent) mit der Gruppe mit eher mehr Stunden (höherfrequent) ergibt deutliche Vorteile für letztere im Zeitverlauf des zweiten Behandlungsjahres in den körperlich akzentuierten Beschwerden, Erschöpfungsneigung, Gliederschmerzen, Herzbeschwerden, Beschwerdedruck und Hypochondrie.

Um eine nach dem ersten Behandlungsjahr erzielte Verbesserung in den psychosomatischen Symptome weiterzuführen, reicht eine niedrigfrequente Behandlung nicht mehr aus, sondern ist offensichtlich eine höhere Sitzungshäufigkeit notwendig.

Dagegen scheint im dritten Behandlungsjahr die Entwicklung der Patienten zu größerer seelischer Gesundheit nicht mehr von der Frequenz der Sitzungen pro Woche abzuhängen.

### *Therapeutengeschlecht*

Die über das Therapeutengeschlecht empirisch-statistisch gebildeten Patientensubgruppen sind zu Anfang der Untersuchung relativ gleich stark mit Krankheitssymptomen belastet, unterscheiden sich also diesbezüglich nicht signifikant voneinander. Dies spricht dafür, dass die Patientenverteilung auf die Therapeuten im statistischen Sinne eher zufällig vonstatten gegangen sein muss und wahrscheinlich keine systematisch wirkende unbekannte Variable den Zuweisungsmodus beeinflusste. Die relative Gleichheit der Subgruppen blieb sowohl im ersten Behandlungsjahr wie auch über den darauf folgenden zweijährigen Behandlungszeitraum erhalten, was bedeutet, dass sich keine relevant unterschiedlichen Therapieeffekte in den Subgruppen ergaben. Man kann daraus schlussfolgern, dass sich die Variable „Therapeutengeschlecht“ in unseren Subgruppen weder begünstigend noch hemmend auf den Therapieerfolg auswirkte.

### *Therapeutenberuf*

Der Vergleich des über die klinischen Skalen gemessenen sich entwickelnden Krankheitszustands der beiden Subgruppen, die entweder bei einem ärztlichen oder bei einem psychologischen Psychotherapeuten in Behandlung waren, ergab weder zu Anfang der Behandlung, noch nach dem ersten Behandlungsjahr und auch nicht nach dem darauf folgenden Behandlungszeitraum von zwei Jahren einen signifikanten Unterschied zwischen den beiden Subgruppen. Dies spricht sowohl dafür, daß sich die Verteilung der

Patienten auf ärztliche oder psychologische Psychotherapeuten in unserer Untersuchungsgruppe offensichtlich weitgehend zufallsbedingt vollzog, als auch dafür, daß sich das Therapeutengeschlecht, zumindest in unserer Untersuchung, nicht bestimmend auf das Ausmaß des Erfolges der Behandlungen ausgewirkt hat.

#### *Schulabschluss*

Die Aufteilung der Patienten unter dem Aspekt des in der Bundesrepublik vorhandenen dreistufigen Schulabschlusssystems führte zum Messzeitpunkt t1 weder zu signifikant unterschiedlich krankheitswertigen Subgruppen und im weiteren Verlauf auch nicht zu unterschiedlich erfolgreichen Therapieverläufen. Ein Einfluss der Variable „Schulbildung der Patienten“ auf den Therapieerfolg konnte zu keinem Zeitpunkt der über drei Jahre hinweg beobachteten Therapieverläufe nachgewiesen werden.

#### *Lie-Verhalten im MMPI(K)*

Die Bildung zweier unterschiedlich therapiewiderständiger Patientensubgruppen mit Hilfe der Lie-Skala des MMPI(K) erzeugte keine klinisch unterschiedlichen Subgruppen: beide mittels der Lie-Skala gebildeten Subgruppen stimmten zum Messzeitpunkt t1 in den klinischen Skalen weitgehend überein, waren also gewissermaßen 'gleich krank'. Beide Patientensubgruppen zeigten zwar hochsignifikante Symptomverbesserungen im Zeitverlauf der Therapie, es konnte jedoch kein signifikanter Verlaufsunterschied in der Entwicklung der beiden Subgruppen festgestellt werden.

#### *Anfänglicher Beschwerdedruck*

Der Mittelwertvergleich der mit Hilfe des Kriteriums der im GBB gemessenen unterschiedlichen Höhe des Beschwerdedrucks gebildeten Subgruppen zeigt, dass die Subgruppe „höherer Beschwerdedruck“ sich auch in fast allen anderen klinischen Skalen „hochsignifikant“ durch höhere Symptomwerte von der Gruppe mit dem „niedrigeren Beschwerdedruck“ unterscheidet. In beiden Patientengruppen können „hochsignifikante“ therapeutische Effekte nachgewiesen werden, die sich hauptsächlich im ersten Behandlungsjahr und - wenngleich in geringerem Umfang - auch noch in dem darauf folgenden Behandlungszeitraum von zwei Jahren ereignen. Bei diesem Verlauf besteht jedoch ein signifikanter Unterschied zwischen den beiden Subgruppen: die „krankere“ Subgruppe mit dem „höheren Beschwerdedruck“ kann bedeutend größere Therapieerfolge für sich verbuchen als die „gesündere“ Gruppe mit dem „geringeren Beschwerdedruck“, wobei sich diese Erfolge wiederum schwerpunktmäßig im ersten und mit einem deutlich geringeren Umfang dann auch noch im darauf folgenden zweijährigen Behandlungszeitraum feststellen lassen.

#### *Anfangsdepression*

Die Subgruppe der Patienten mit „höherer Anfangsdepression“ zeigt auch in den anderen klinischen Skalen signifikant höhere Mittelwerte als die Subgruppe mit „geringerer Anfangsdepression“, und dies sowohl zu Beginn wie auch am Ende aller beobachteten Behandlungsperioden. Die hochsignifikanten Therapieeffekte vollziehen sich schwerpunktmäßig im ersten Jahr der Behandlung und nur in einem geringeren Umfang im darauf folgenden Zweijahreszeitraum. Dabei ereignen sich diese Therapieeffekte in einem

weit größeren Umfang in der Subgruppe mit der „höheren Anfangsdepression“ als in der Subgruppe mit der „geringeren Anfangsdepression“.

## **11. Die Katamneseuntersuchung**

### **11.1 Einleitung**

#### **11.1.1 Zum Problem der Katamnese in der Psychoanalyse**

In der psychoanalytischen Forschungstradition bestand lange Zeit wenig Interesse an katamnestischen Untersuchungen der eigenen Arbeit.

Der Widerstand gegen Katamnesen in der psychoanalytischen Forschung hat mit Vorbehalten gegenüber Forschung in der Psychoanalyse überhaupt zu tun (Kernberg, 1997: z.B. mit der Begründung ihrer Komplexität, der Simplizität empirischer Methoden u.a.), aber auch mit dem Verständnis vom Ende der analytischen Behandlung, bei dem idealerweise die Übertragung aufgelöst sein sollte und ein nachträglicher postanalytischer Kontakt nur eine unaufgelöste Übertragungs-Gegenübertragungs-Beziehung widerspiegeln würde (Wallerstein, 1997). Demgegenüber stand Freud katamnestischen Informationen nicht abgeneigt gegenüber, er war im Gegenteil froh, wenn er sich ein Bild über den Gesundheitszustand ehemaliger Patienten machen konnte.

Wir referieren hier kurz über zwei wichtige Forschungsprojekte, die sich mit dem Thema der Übertragung in der Katamneseuntersuchung befassen:

1) Pfeffer, 1959 (zit. bei Bachrach et al., 1991), gewann seine Daten in psychoanalytischen Interviews, die von einem sog. Katamnese-Analytiker zu einem Katamnesezeitpunkt ungefähr vier Jahre nach Abschluss der Behandlung durchgeführt wurden. Man nahm an, dass der beste Weg, psychoanalytische Daten zum psychoanalytischen Prozess zu gewinnen, die möglichst genaue Annäherung an die vorhergegangene psychoanalytische Situation sei. Es wurden durchschnittlich fünf bis sechs Gespräche geführt, mit dem einheitlichen Ergebnis eines (gewünschten) Wiederaufflackerns der zentralen Übertragungen aus der damaligen Analyse, die es möglich machten, den damaligen Analyseverlauf zu rekonstruieren.

2) In der Katamnese studie der Psychotherapie-Forschungsgruppe der Menninger Foundation (Wallerstein, 1997) wurden ebenfalls Interviews geführt, jedoch nicht, wie bei Pfeffer, in Form einer Mini-Analyse mit neuerlichem Übertragungsgeschehen, sondern strukturierter, um ein Übertragungsaufleben zu vermeiden. Außerdem wurden projektive Tests verwendet und indirekte Quellen (Verwandte) genutzt. Der Katamnesezeitpunkt lag zwischen zwei und drei Jahren nach Behandlungsende, obwohl fünf Jahre als optimal angesehen wurden. Die Studie zeigt, dass systematische Katamnesen unterschiedliche Auswirkungen auf die Beendigungsphase der Psychoanalysen haben: manchmal komplizieren sie die Beendigung, manchmal bewirkt die geplante Katamnese auch eine günstigere Gestaltung des Abschlusses der Analyse und der Ablösung vom ehemaligen Analytiker.

#### **11.1. 2 Das Design unserer Katamnese studie**

Ursprünglich enthielt unser Studiendesign keine katamnestische Untersuchung. Um die Effektivität von individualpsychologisch-psychoanalytischer Psychotherapie in Einzelsitzungen überprüfen zu können,

erschien uns bei der Planung dieser Untersuchung eine Katamnese nicht angemessen zu sein. Unter Psychoanalytikern galt es oft als Verstoß gegen die Regeln der psychoanalytischen Behandlung, „den ehemaligen Analysanden zum Objekt katamnestischer Studien zu machen“ (Leuzinger-Bohleber, 1997, S.11). Unsere Auffassung änderte sich im Laufe der Zeit, als immer häufiger die Erfordernis des langfristigen Wirksamkeitsnachweises von Psychotherapie nach Behandlungsende betont wurde (so z.B. im Heidelberger Katamneseprojekt).

Es stellte sich für uns nun die Frage, wie wir die Katamnese konzipieren sollten. In der Logik der vorangegangenen Effektivitätsstudie hätte es sich natürlich angeboten, die Testbatterie der Verlaufsstudie den ehemaligen Patienten nochmals vorzulegen, um im direkten Vergleich der Ergebnisse die weitere Entwicklung nach Behandlungsende feststellen zu können. Wir haben jedoch zugunsten von mehr qualitativen Verfahren darauf verzichtet, die es uns erlauben sollten, mehr über die subjektive Einschätzung von Therapie und Therapieerfolg der ehemaligen Patienten und ihrer Haltung zu sich selbst nach Beendigung ihrer Behandlung zu erfahren. Dabei sollte auch eine gewisse Vergleichbarkeit mit anderen Studien gewährleistet sein.

Folgende Verfahren (genaue Beschreibung s. u. Kap. 11.2.3 u. 11.2.4) wurden von uns eingesetzt: der Helping Alliance Questionnaire (Luborsky), der Intrex-Fragebogen (Benjamin, Tress) und drei Fragen zur Freude in der Therapie.

Entfallen musste die Fortschreibung der Testergebnisse der Behandlungszeit in den Katamnesezeitraum ebenso wie eine mit Hilfe des Intrex-Fragebogens mögliche Untersuchung der Einschätzung der Beziehung zu den wichtigen Personen der Vergangenheit und Gegenwart, um daran den Zustand psychischer Gesundheit nach dem Behandlungsende feststellen zu können. Dies hätte den Rahmen der schriftlichen Befragung gesprengt und wäre für die Patienten auch nicht zumutbar gewesen.

## **11.2 Auswertung der Katamnesedaten**

Die angefallenen Katamnesedaten haben wir unter folgenden Gesichtspunkten ausgewertet:

Wer sind die Katamneseteilnehmer?

Wie viele stammen aus der Untersuchungs-, wie viele aus der Kontrollgruppe?

Wie lange liegt das Therapieende zurück und gibt es Unterschiede zwischen den Teilnehmern, deren Therapieende länger oder kürzer zurückliegt?

Welche soziodemografische Merkmale kennzeichnen die Katamneseteilnehmer?

Wir überprüfen dann, ob sich die Katamneseteilnehmer von den Nichtteilnehmern unterscheiden und ziehen dazu die Merkmale Alter, Geschlecht und Globalmaß der Therapieeffektivität heran.

Zwei Effektivitätsmaße in der Katamneseuntersuchung sollen Aussagen über die nachhaltige Wirksamkeit der psychotherapeutischen Behandlung erlauben:

1) Die direkte globale Einschätzung von Veränderungen durch die Patienten selbst mit Hilfe des Helping Alliance Questionnaire-HAQ von Luborsky.

2) Das theoretisch begründete katamnestiche Ergebnismaß - das Intrex-Introjekt - im Vergleich zu anderen Untersuchungen.

Danach schauen wir uns die Antworten auf drei Fragen zur Freude in der Psychotherapie an.

Zum Schluss berechnen wir einige Korrelationen, um Zusammenhänge zwischen verschiedenen Katamneseergebnissen sowie zwischen Katamneseergebnissen und Testergebnissen aus der Untersuchungsgruppe festzustellen.

### 11.2.1 Wer sind die Katamneseteilnehmer?

#### 11.2.1.1 Patienten der Untersuchungsgruppe und der Kontrollgruppe

**Tabellen 79: Katamneseteilnehmer aus UG und KG**

	<b>N</b>	<b>%</b>
<b>teilgenommen:</b>	57	100,0
<b>aus UG</b>	45	78,9
<b>aus KG</b>	12	21,1

	<b>N</b>		<b>%</b>	
	<b>UG (t1)</b>	<b>KG (t1)</b>	<b>UG (t1)</b>	<b>KG (t1)</b>
<b>teilgenommen</b>	45	13	31,0	25,5
<b>nicht teilgenommen</b>	100	38	69,0	74,5
<b>Gesamt</b>	145	51	100,0	100,0

Tabelle 79 zeigt, dass sich die Katamnesegruppe zu über 3/4 aus Patienten der Untersuchungsgruppe und zu knapp 1/4 der Kontrollgruppe zusammensetzt. 31% der zu t1 in der Untersuchungsgruppe befindlichen 145 Patienten haben an der Katamneseuntersuchung teilgenommen, 25 % der zu t1 bestehenden Kontrollgruppe von 51 Patienten.



### 11.2.1.2 Der Zeitpunkt ihres Therapieendes

**Tabelle 80: Katamneseteilnehmer; Monat und Jahr ihres Therapieendes, Häufigkeiten und Prozente**

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
8/94	1	1,8	2,2	2,2
11/94	1	1,8	2,2	4,4
4/95	1	1,8	2,2	6,7
7/95	1	1,8	2,2	8,9
12/95	1	1,8	2,2	11,1
3/96	1	1,8	2,2	13,3
7/96	1	1,8	2,2	15,6
3/97	1	1,8	2,2	17,8
4/97	2	3,5	4,4	22,2
5/97	1	1,8	2,2	24,4
6/97	1	1,8	2,2	26,6
7/97	1	1,8	2,2	28,8
8/97	1	1,8	2,2	31,1
9/97	1	1,8	2,2	33,3
11/97	1	1,8	2,2	35,6
12/97	3	5,3	6,7	42,2
2/98	1	1,8	2,2	44,4
3/98	1	1,8	2,2	46,7
4/98	2	3,5	4,4	51,1
5/98	1	1,8	2,2	53,3
6/98	4	7,0	8,9	62,2
8/98	1	1,8	2,2	64,4
9/98	1	1,8	2,2	66,7
11/98	1	1,8	2,2	68,9
12/98	2	3,5	4,4	73,3
1/99	1	1,8	2,2	75,6
2/99	2	3,5	4,4	80,0
4/99	1	1,8	2,2	82,2
5/99	1	1,8	2,2	84,4
9/99	1	1,8	2,2	86,7
10/99	1	1,8	2,2	88,9
11/99	1	1,8	2,2	91,1
1/00	3	5,3	6,7	97,8
2/00	1	1,8	2,2	100,0
Gesamt	45	78,9	100,0	
Fehlend	12	21,1		
Gesamt	57	100,0		

**Tabelle 81: Katamneseteilnehmer:  
Therapieende länger und kürzer  
zurückliegend  
(Trennung über dem Median)**

Therapieende

Zeitraum	N	%
8/94-3/98	23	51,1
4/98-2/00	22	48,9
Gesamt	45	100,0

Der Zeitpunkt des Therapieendes liegt im Zeitraum 8/94 bis 2/00, der Median liegt bei 4/98.

Bei der Hälfte der Katamneseteilnehmer liegt demnach das Therapieende, vom Katamnesezeitpunkt aus gerechnet, zwei bis sechs Jahre zurück, bei der anderen Hälfte liegt das Therapieende im Zeitraum „gerade bis vor weniger als zwei Jahren beendet“. Das Gros der Katamneseteilnehmer hat die Therapie in den Jahren 1997 bis 1998 beendet (55 %) (siehe Tabelle 80). Der Zeitraum zwischen Therapieende und Katamnesezeitpunkt ist in vielen Studien unterschiedlich lang (z. B.: Heidelberger Katamneseprojekt durchschnittlich 3,5 Jahre, Menninger-Studie zwischen zwei u. drei Jahren).

Smith et al. (1980) fanden zwischen der Effektstärke und dem Nachuntersuchungszeitraum bis zu zwei Jahren eine starke negative Korrelation. Auch Kazdin (1994) zeigte, dass sowohl die absoluten als auch die relativen Effekte der Behandlung im Verlauf der Zeit schwanken können. Shapiro et al. (1995) zeigte, dass man bei der Nachuntersuchung nach drei Monaten bzw. nach einem Jahr zu ganz unterschiedlichen Einschätzungen der Langzeitergebnisse kommt, je nachdem, in welcher Form und wie lange Therapien durchgeführt wurden.

#### **11.2.1.3 Soziodemografische Merkmale der Katamneseteilnehmer**

Siehe Tabellen 83 a - f im Anhang

Die Verteilung nach Geschlecht ist nicht anders als in der Untersuchungsgruppe zum Zeitpunkt t1, während sich in der Altersverteilung der Zeitabstand zwischen Untersuchungsbeginn und Katamnesezeitpunkt zeigt: Die Patienten sind einige Jahre älter geworden, entsprechend ist die Altersgruppe 21-30 Jahre nur noch mit einem Patienten vertreten und über 90 % befinden sich in der zusammengefassten Altersgruppe 31-50 Jahre.

Der Vergleich mit der Bevölkerungsstatistik (Tabelle 82) macht deutlich, dass in der Gruppe der Katamneseteilnehmer die beiden Altersgruppen 21-30 und 51-60 Jahre stark unter-, während die Altersgruppen 31-40 und 41-50 überrepräsentiert sind.

Vom Familienstand her betrachtet zeigt sich, dass die Mehrheit verheiratet ist (fast die Hälfte), fast 1/3 ist ledig, ungefähr 1/5 ist getrennt lebend oder geschieden.

2/5 der Katamneseteilnehmer ist kinderlos; wer Kinder hat, hat meistens zwei. Wenige Teilnehmer haben ein Kind, drei oder vier Kinder.

Weit über die Hälfte haben als Schulabschluss das Abitur, über ein Viertel die mittlere Reife. Entsprechend dem niedrigen Anteil des Hauptschulabschlusses ist auch der Anteil der Facharbeiter sehr gering. Bei den Berufen überwiegen die Beamten und Angestellten vor den Selbstständigen.

Es besteht eine große Übereinstimmung mit der Untersuchungsgruppe (vgl. Tabelle 7.4).

## 11.2.2 Unterscheiden sich die Teilnehmer der Katamnese von den Nichtteilnehmern?

### 11.2.2.1 Altersverteilung der Katamneseteilnehmer vs. Nichtteilnehmer

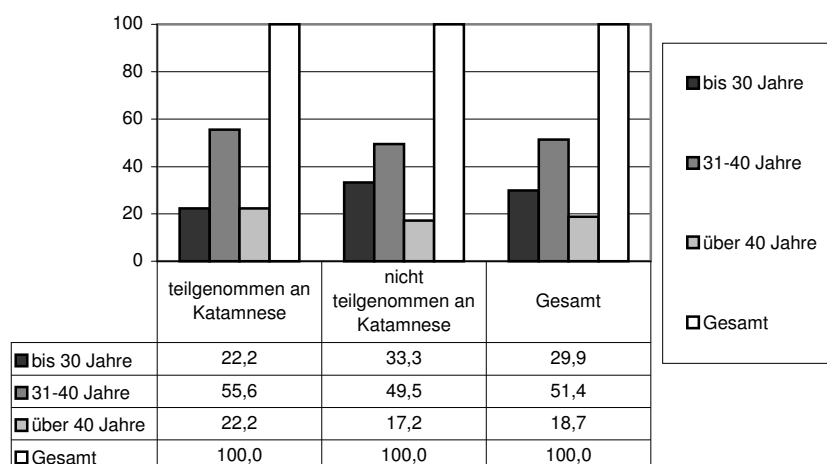
**Tabelle 84: Altersverteilung\* der Katamnese-Teilnehmer vs. Nichtteilnehmer aus der UG**

	bis 30		31-40		über 40		Gesamt	
	N	%	N	%	N	%	N	%
teilgenommen an Katamnese	10	22,2	25	55,6	10	22,2	45	100,0
nicht teilgenommen an Katamnese	33	33,3	49	49,5	17	17,2	99	100,0
Gesamt	43	29,9	74	51,4	27	18,7	144	100,0

\* Alter zu t1

Tabelle 84 macht deutlich, dass die Teilnehmer an der Katamnese tendenziell älter sind als die Nichtteilnehmer, der Prozentanteil der 30 bis 40Jährigen und der über 40Jährigen ist bei den Teilnehmern größer als bei den Nichtteilnehmern, während in der Gruppe der bis 30Jährigen der Anteil der Nichtteilnehmer größer ist als der der Teilnehmer (Abb. 25).

**Abb. 25: Altersverteilung (in Prozent) der Katamneseteilnehmer und Nichtteilnehmer aus der Untersuchungsgruppe**



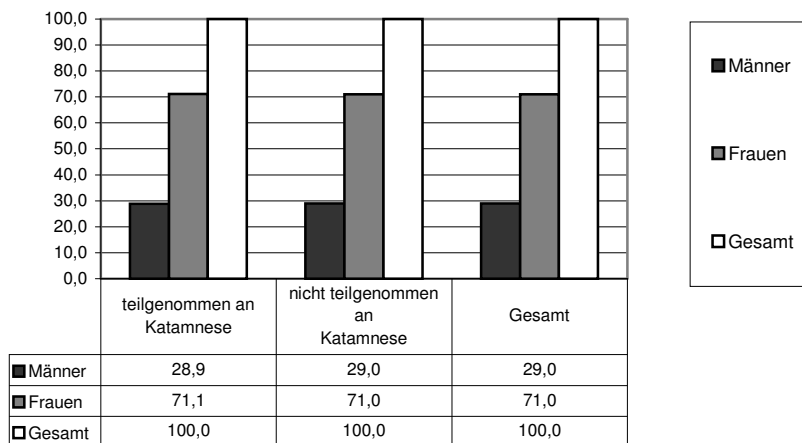
### 11.2.2.2 Geschlechterverteilung der Katamneseteilnehmer vs. Nichtteilnehmer

**Tabelle 85: Geschlechterverteilung der Katamneseteilnehmer vs. Nichtteilnehmer aus der UG**

	Männer		Frauen		Gesamt	
	N	%	N	%	N	%
teilgenommen an Katamnese	13	28,9	32	71,1	45	100,0
nicht teilgenommen an Katamnese	29	29,0	71	71,0	100	100,0
<b>Gesamt</b>	<b>42</b>	<b>29,0</b>	<b>103</b>	<b>71,0</b>	<b>145</b>	<b>100,0</b>

Tabelle 85 zeigt: Männer und Frauen sind in beiden Gruppen gleich verteilt, Katamneseteilnehmer und Nichtteilnehmer unterscheiden sich nicht nach Geschlecht (s. auch Abb. 26).

**Abb. 26: Geschlechterverteilung (in Prozent) der Katamneseteilnehmer und Nichtteilnehmer aus der Untersuchungsgruppe**



### 11.2.2.3 Globalmaß der Therapieeffektivität (Effektstärke t1:t5) für Katamneseteilnehmer und Nichtteilnehmer

**Tabelle 86: Therapieeffekt (t1:t5) der Katamneseteilnehmer vs. Nichtteilnehmer aus der UG**

An Katamnese		Effekt t1:t5
teilgenommen (N=26)	M s	0,822 0,9159
nicht teilgenommen (N=31)	M s	0,8165 1,0015
insgesamt (N=57)	M s	0,819 0,9549

Tabelle 86 macht deutlich, dass die Effektstärke (t1:t5) bei den Patienten, die an der Katamnese teilgenommen haben, gleich groß ist wie bei den Patienten, die nicht teilgenommen haben. Damit kann ausgeschlossen werden, dass an der Katamnese nur die erfolgreicheren Patienten teilgenommen haben und damit das Katamneseergebnis einseitig positiv verfälscht wäre.

Insgesamt wird deutlich, dass sich Teilnehmer und Nichtteilnehmer an der Katamneseuntersuchung in den untersuchten Merkmalen kaum (außer im Alter) voneinander unterscheiden.

### 11.2.3 Zwei verschiedene Effektivitätsmaße für die Katamnese

Mit zwei verschiedenen Effektivitätsmaßen wurde die nachhaltige Wirksamkeit der psychotherapeutischen Behandlung nach ihrer Beendigung überprüft.

Zur direkten globalen Einschätzung von Veränderungen durch die Patienten selbst innerhalb der Katamneseuntersuchung haben wir den HAQ eingesetzt.

#### 11.2.3.1 Der Helping Alliance Questionnaire (HAQ)

(Der Fragebogen befindet sich im Anhang Kapitel 14.)

Dieser, von Luborsky (1984, zit. bei Bassler, 1995a) entwickelte Fragebogen, wurde von Bassler u.a. (1995a) ins Deutsche übersetzt, um seine Verwendbarkeit im Rahmen von stationärer Psychotherapie zu überprüfen. Der Fragebogen besteht aus zwölf Fragen zur therapeutischen Beziehung und Prozessvariablen und kann sowohl vom Patienten als auch vom Therapeuten beantwortet werden. Analysen dieses Fragebogens (patientenseitige Einschätzung) ergaben Hinweise auf zwei immanente Faktoren: „therapeutische Beziehungszufriedenheit“ und „Zufriedenheit mit dem therapeutischen Ergebnis“.

Zusammenfassend halten die Autoren diesen Fragebogen für ein interessantes und vielversprechendes Instrument zur Evaluation des therapeutischen Prozesses, das leicht in der klinischen Routine eingesetzt werden könne.

Beide Subskalen des HAQ erweisen sich als ausreichend variabel und valide und erlauben deutlich zwischen erfolgreichen und weniger erfolgreichen Patienten zu unterscheiden. Dabei scheinen beide Skalen trotz der gemeinsamen aufgeklärten Varianz jeweils spezifische Aspekte in der patientenseitigen Bewertung zu erfassen, wie die differentiellen Ergebnisse in Bezug auf unabhängig erhobene Maße zeigen.

Spez. die Subskala „Erfolgszufriedenheit“ korrelierte vergleichsweise hoch mit dem unabhängigen kombinierten Erfolgskriterium (z.B. Skala Beschwerdedruck des GBB) (Bassler, 1995 S. 31). Patienten, die bezüglich ihres erreichten Therapieerfolgs zufriedener bzw. optimistischer sind, fühlen sich offenbar auch verändert hinsichtlich beeinträchtigender Körpersymptome. Diese Veränderungen lassen sich mit dem subjektiv erlebten Therapieerfolg in Zusammenhang bringen, womit allerdings nicht die Frage nach Ursache und Wirkung beantwortet ist. Die Skala „Erfolgszufriedenheit“ kann als Indikator für den allgemeinen Therapieerfolg gelten. Aus der Sicht der Autoren kann der HAQ damit sowohl als Prozessforschungs- als auch als Outcomemaß eingesetzt werden, wobei die „Erfolgszufriedenheit“ eher Outcomequalitäten besitzt und sich die „Beziehungszufriedenheit“ eher als Prozessforschungsmaß empfiehlt.

Wir haben uns beim Einsatz des HAQ leiten lassen von der Empfehlung Orlinsky & Howards (1986), patientenseitige Daten zur Erfolgsmessung Therapeutendaten vorzuziehen, um auf jeden Fall die Erwartungshaltungen der Patienten und nicht stattdessen oder zusätzlich die der Therapeuten zu berücksichtigen. Dies wäre auch aus forschungsökonomischen Erwägungen heraus nur unter Schwierigkeiten möglich gewesen.

Die Aussagen 1, 2, 6, 7, 8, 9, 10 wurden von uns vom Präsens ins Perfekt bzw. Imperfekt gesetzt, um der Tatsache Rechnung zu tragen, dass das Behandlungsende zum Katamnesezeitpunkt bis zu sechs Jahre zurücklag.

Außerdem haben wir zur einfacheren Handhabung für die Patienten auch für Aussage 12 die sechsstufige Skalierung (statt sieben im Original) verwendet und Schulnoten (von 1 bis 6) vergeben lassen (statt 1 = sehr viel verschlechtert bis 7 = sehr viel verbessert, im Original).

**Tabelle 87: Der Helping Alliance Questionnaire (HAQ) in der von uns verwendeten, leicht veränderten Form**

1. Ich glaube, dass mein Therapeut mir geholfen hat.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	1	1,8	1,8
2	3	5,3	7,0
3	4	7,0	14,0
4	4	7,0	21,1
5	25	43,9	64,9
6	20	35,1	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.

6 = Ich halte das für sehr zutreffend.

2. Ich glaube, dass die Behandlung mir geholfen hat.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	2	3,5	3,5
2	1	1,8	5,3
3	4	7,0	12,3
4	8	14,0	26,3
5	20	35,1	61,4
6	22	38,6	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.

6 = Ich halte das für sehr zutreffend.

Tabelle 87 (Fortsetzung)

3. Ich habe einige neue Einsichten gewonnen.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	2	3,5	3,5
2	1	1,8	5,3
3	2	3,5	8,8
4	2	3,5	12,3
5	15	26,3	38,6
6	35	61,4	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.  
6 = Ich halte das für sehr zutreffend.

4. Seit einiger Zeit fühle ich mich besser.

	Häufigkeit	Prozent	Kumulierte Prozente
1	2	3,5	3,5
2	2	3,5	7,0
3	5	8,8	15,8
4	13	22,8	38,6
5	21	36,8	75,4
6	14	24,6	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.  
6 = Ich halte das für sehr zutreffend.



Tabelle 87 (Fortsetzung)

5. Ich sehe, dass ich die Probleme bewältigen kann, wegen derer ich zur Behandlung kam.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	4	7,0	7,0
2	1	1,8	8,8
3	3	5,3	14,0
4	16	28,1	42,1
5	19	33,1	75,4
6	14	24,6	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.

6 = Ich halte das für sehr zutreffend.

6. Ich hatte das Gefühl, mich auf den Therapeuten verlassen zu können.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	3	5,3	5,3
2	2	3,5	8,8
3	1	1,8	10,5
4	7	12,3	22,8
5	13	22,8	45,6
6	31	54,4	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.

6 = Ich halte das für sehr zutreffend.

Tabelle 87 (Fortsetzung)

7. Ich hatte das Gefühl, dass mich der Therapeut versteht.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	1	1,8	1,8
2	2	3,5	5,3
3	2	3,5	8,8
4	11	19,3	28,1
5	16	28,1	56,1
6	25	43,9	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.

6 = Ich halte das für sehr zutreffend.

8. Ich hatte das Gefühl, dass mein Therapeut möchte, dass ich meine Ziele erreiche.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	3	5,3	5,3
2	2	3,5	8,8
3	1	1,8	10,5
4	10	17,5	28,1
5	15	26,3	54,4
6	26	45,6	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.

6 = Ich halte das für sehr zutreffend.

Tabelle 87 (Fortsetzung)

9. Ich hatte das Gefühl, dass ich wie auch der Therapeut ernsthaft an einem Strang ziehen.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	1	1,8	1,8
2	3	5,3	7,0
3	5	8,8	15,8
4	8	14,0	29,8
5	13	22,8	52,6
6	27	47,4	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.

6 = Ich halte das für sehr zutreffend.

10. Ich glaube, dass ich und der Therapeut meine Probleme ähnlich gesehen und beurteilt haben.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	1	1,8	1,8
2	4	7,0	8,9
3	8	14,0	23,2
4	14	24,6	48,2
5	16	28,1	76,8
6	13	22,8	100,0
Gesamt	56	98,2	
Fehlend	1	1,8	
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.

6 = Ich halte das für sehr zutreffend.

Tabelle 87 (Fortsetzung)

11. Ich habe das Gefühl, dass ich mich jetzt selbst verstehen und mit mir auseinandersetzen kann.

	Häufigkeit	Prozent	Kumulierte Prozente
1*	2	3,5	3,5
3	2	3,5	7,0
4	11	19,3	26,3
5	23	40,4	66,7
6	19	33,3	100,0
Gesamt	57	100,0	

\* 1 = Ich halte das für sehr unzutreffend.

6 = Ich halte das für sehr zutreffend.

12. Insgesamt schätze ich den Therapieerfolg ein als:

	Häufigkeit	Prozent	Kumulierte Prozente
1*	14	24,6	24,6
2	22	38,6	63,2
3	13	22,8	86,0
4	4	7,0	93,0
5	4	7,0	100,0
6	0	0,0	100,0
Gesamt	57	100,0	

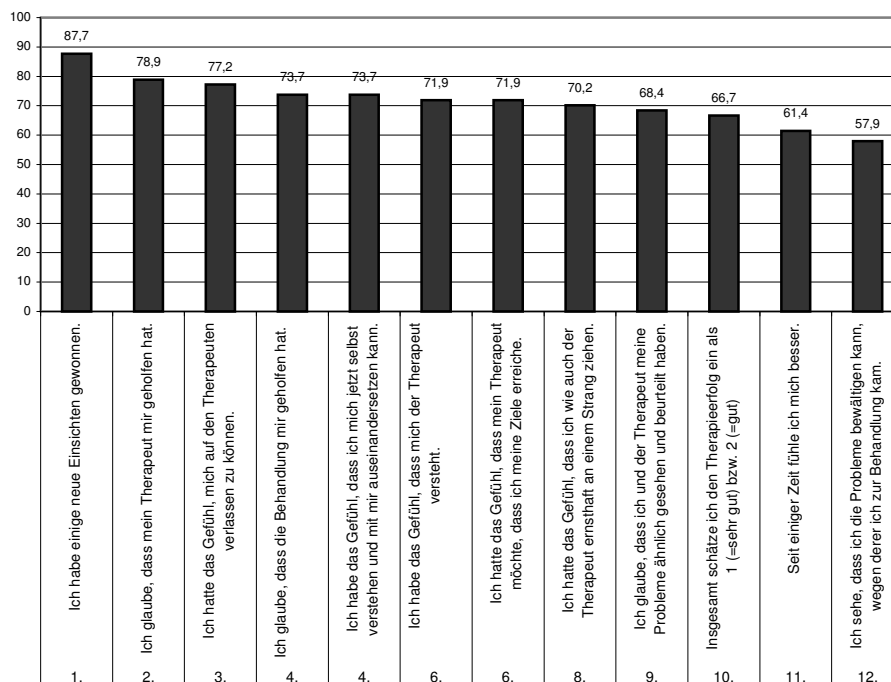
\* 1 = sehr gut

6 = sehr schlecht

Bringt man die Fragen des HAQ in eine Rangreihe und legt dabei die Häufigkeit der Zustimmung mit 5 + 6 (= ich halte das für sehr zutreffend) zugrunde, so ergibt sich folgende Reihenfolge, beginnend mit dem höchsten Grad der Zustimmung:

1. Ich habe einige neue Einsichten gewonnen. (87,7%)
2. Ich glaube, dass mein Therapeut mir geholfen hat. (78,9%)
3. Ich hatte das Gefühl, mich auf den Therapeuten verlassen zu können. (77,2%)
4. Ich glaube, dass die Behandlung mir geholfen hat. (73,7%)  
Ich habe das Gefühl, dass ich mich jetzt selbst verstehen und mit mir auseinandersetzen kann. (73,7%)
6. Ich hatte das Gefühl, dass mich der Therapeut versteht. (71,9%)  
Ich hatte das Gefühl, dass mein Therapeut möchte, dass ich meine Ziele erreiche. (71,9%)
8. Ich hatte das Gefühl, dass ich wie auch der Therapeut ernsthaft an einem Strang ziehen. (70,2%)
9. Ich glaube, dass ich und der Therapeut meine Probleme ähnlich gesehen und beurteilt haben. (68,4%)
10. Insgesamt schätze ich den Therapieerfolg ein als 1 (= sehr gut) bzw. 2 (= gut). (66,7%)
11. Seit einiger Zeit fühle ich mich besser. (61,4%)
12. Ich sehe, dass ich die Probleme bewältigen kann, wegen derer ich zur Behandlung kam. (57,9%)

**Abb. 27: Rangreihe (in Prozent) der HAQ-Items nach dem Grad der Zustimmung**



Es zeigt sich, dass die Einsichtsgewinnung durch die psychotherapeutische Behandlung am uneingeschränktesten betont wird, während die Problembewältigung (item 12 ) von allen statements die geringste Zustimmung (aber immerhin noch von nahezu 60% der Katamneseteilnehmer) erfahren hat (s. auch Abb. 27). Dies bestätigt vielleicht Befunde, die mehr die einsichtsfördernde Wirkung der psychoanalytischen Behandlung und weniger die Symptom-(= Problem-)beseitigung zeigen (siehe z.B. v. Rad et al.,1998).

**Tabelle 88: Mittelwerte und Standardabweichungen der beiden Faktoren „therapeutische Beziehung“ und „therapeutischer Erfolg“ des Helping Alliance Questionnaire (HAQ)**

	<b>M</b>	<b>s</b>
<b>HAQ-Beziehung</b>	29,3334	6,2202
<b>HAQ-Erfolg</b>	28,7892	5,6814

Tabelle 88 zeigt, dass sich die Mittelwerte der beiden Faktoren „therapeutische Beziehung“ und „therapeutischer Erfolg“ nur unwesentlich unterscheiden, der Mittelwert für den Faktor HAQ-Beziehung ist etwas größer als der für HAQ-Erfolg.

Die Behandlungszufriedenheit unserer Patienten kann demnach als hoch eingestuft werden.

„Behandlungszufriedenheit“ ist nach Rudolf (1994, S. 31) insbesondere bei den ambulanten psychoanalytisch orientierten Therapien sehr hoch. Die entsprechenden Einschätzungen bei den prognostisch schwierigeren stationären Patienten fallen nach Rudolf ungünstiger aus.

**Tabelle 89: Mittelwerte und Standardabweichungen der beiden Faktoren „therapeutische Beziehung“ und „therapeutischer Erfolg“ des Helping Alliance Questionnaire (HAQ); Vergleich mit den Ergebnissen von Kammerer (1998) und Bassler et al. (1993. zit.b. Kammerer. 1998)**

	therapeutische Beziehung		therapeutischer Erfolg	
	<b>M</b>	<b>s</b>	<b>M</b>	<b>s</b>
unsere Studie	29,3	6,2	28,8	5,7
Kammerer	29,8	4,6	17,4	4,2
Bassler et al.	30,3	4,1	16,6	4,2

Tabelle 89 macht deutlich, dass sich die Mittelwerte der drei Studien für den Faktor „therapeutische Beziehung“ kaum unterscheiden, während der Mittelwert für den Faktor „therapeutischer Erfolg“ in unserer Studie größer ist als in den beiden anderen Studien, deren Mittelwerte wiederum fast gleich groß sind. Der therapeutische Erfolg wird von den Patienten unserer Studie größer eingeschätzt als von den Patienten der beiden anderen Studien, was mit den unterschiedlichen Stichproben (Klinikpatienten), aber auch mit der längeren Behandlungszeit unserer Patienten zusammenhängen kann.

Zusammenfassung der Ergebnisse des HAQ: Die Behandlungszufriedenheit unserer Patienten ist sehr hoch, sowohl was die therapeutische Beziehung als auch den therapeutischen Erfolg angeht. Die Einsichtsgewinnung durch die Behandlung wird von den Patienten uneingeschränkt betont.

### 11.2.3.2 Die Intrex-Kurzform

(Der Fragebogen befindet sich im Anhang Kapitel 14.)

Neben dem HAQ und seiner direkten Einschätzung von Veränderungen durch die Patienten verwendeten wir auch ein theoretisch begründetes Ergebnismaß – das Intrex-Introjekt – und setzten es in Vergleich mit anderen Untersuchungen.

(Bei der Darstellung des Intrex-Fragebogens und seiner Auswertung orientieren wir uns vorrangig an Tress, 1993 und Kammerer, 1998).

Zunächst eine Einführung in den Intrex-Fragebogen, die ihm zugrundeliegende Theorie und seine Auswertung.

Der Intrex-Fragebogen basiert auf dem SASB-Modell (Strukturelle Analyse Sozialen Verhaltens).

Die Leitfrage von SASB lautet (Tress, 1993, S. 35): „Wer kommuniziert qualitativ wie, mit wem und wie reagiert jene(r) darauf?“

SASB baut auf der Persönlichkeitstheorie von Sullivan auf, in dem sie einen Zusammenhang zwischenmenschlicher Erfahrungen mit der Entwicklung psychischer Strukturen annimmt. Persönlichkeit wird als verinnerlichtes, relativ überdauerndes Muster sich wiederholender, charakteristischer interpersoneller Situationen gefasst. Die Verinnerlichung zwischenmenschlicher Erfahrungen wird im Konzept der Introjektion gefasst.

SASB liegen die beiden Dimensionen Affiliation und Dominanz zu Grunde. Bei der Analyse wird hinsichtlich des Aufmerksamkeitsfokus der Kommunikation unterschieden:

- a) Transitiv: aktiv auf den Partner gerichtet, den Partner beeinflussend, elternartig.
- b) Intransitiv: reaktiv auf den Sprecher selbst bezogen, Sprecher zeigt „Wirkung“, kindartig.
- c) Introjekt: selbstbezüglicher Kommentar, auf sich selbst gerichtete Handlung, verinnerlichte Elternhaltung.

Nach Benjamin lässt sich jede zwischenmenschliche Verhaltensweise auf die Dimensionen Interdependenz und Affiliation sowie hinsichtlich des Aufmerksamkeitsfokus einordnen. Durch die Kombination der drei genannten Dimensionen ergibt sich die Modellstruktur der SASB Cluster-Version. Die horizontale Achse stellt jeweils die Dimension Affiliation dar und auf der vertikalen Achse wird die Interdependenz abgebildet.

Mit Hilfe standardisierter Verfahren werden die Daten erhoben. Es handelt sich um eine Codiermethode zur Verhaltensbeobachtung und zwei Fragebögen (Intrex und Intrex-Kurzform).

1987 wurde die Intrex-Kurzform von Benjamin eingeführt, die erste deutsche Übersetzung wurde 1989 von Tress und Benjamin vorgenommen; die erste Revision dieser Übersetzung existiert seit Ende 1991. In der Intrex-Kurzform ist für jeden der acht Oktanten des SASB-Modells (s. Abb. 28) – d.h. für jedes „Cluster“ – ein Item formuliert, das inhaltlich den gesamten Bereich des entsprechenden Oktanten abdecken soll. acht Items beschreiben den Umgang mit der eigenen Person („Introjekt“) und 16 Items interpersonelle Transaktionen, die wir aber aus forschungsökonomischen Gründen nicht verwendet haben. Der von uns eingesetzte Fragebogen bezieht nur die Fragen zum Introjekt ein, der Umgang mit sich selbst nach Abschluss der Behandlung sollte im Mittelpunkt unseres Interesses stehen. Dies war möglich, weil die Testautoren ausdrücklich die Abwandlung der Standardserie je nach klinischer oder wissenschaftlicher Fragestellung empfehlen. Dass sich die Introjekte zur Beschreibung therapeutischer Veränderungen heranziehen lassen, wurde auch in der unten angeführten Isnyer Untersuchung festgestellt. Das Introjekt wird in der von uns benutzten Fassung in zwei Versionen erfragt: „in guten Zeiten“ und „in schlechten Zeiten“.

Es liegen Ergebnisse von drei Untersuchungen vor, auf die wir uns, im Vergleich mit unseren Ergebnissen im wesentlichen beziehen:

Die erste Untersuchung stammt von Benjamin, SASB: Short Form User`s Manual, 1988. Die zweite stammt von Davies-Osterkamp, Hartkamp u. Junkert in: Tress (1993). Untersucht wurden Studenten (N=101) und Patienten (N=51). Die dritte Untersuchung wurde 1998 von Kammerer vorgelegt. Sie stammt aus der psychosomatischen Klinik Alpenblick in Isny. Die Stichprobe besteht aus 192 Patienten (= 28 % Beteiligung), die zwischen 1991 und 1993 in dieser Klinik behandelt wurden.

Gütekriterien werden für die Intrex-Kurzform von Benjamin (1988), (nachzulesen bei Davies-Osterkamp et al., 1993) angegeben, wobei sie ausschließlich an studentischen Gruppen („Normalpersonen“) überprüft wurden:

*Retest-Reliabilität* der Cluster-Scores für das Introjekt (überprüft an zwölf Personen), Zeitabstand ein Monat:  $r = .796$  (gute Zeiten),  $r = .655$  (schlechte Zeiten), der Patternkoeffizienten ATK, CON und CFL: zusammengefaßt für alle Beziehungen ATK  $r = .322$  bis  $.956$ , CON  $r = .395$  bis  $.830$  CFL  $r = .085$  bis  $.790$ . Die Beziehungsbeurteilungen in „schlechten Zeiten“ waren insgesamt weniger stabil als „in guten Zeiten“.

Zur *Validität*: Für die Intrex-Kurzform erbrachte die Überprüfung der faktoriellen Struktur, getrennt für die beiden Parallelformen, so die Autoren, befriedigende Resultate. Die ersten beiden Faktoren erklärten zwischen 55 % und 70 % der Varianz, die Items waren in der erwarteten zirkulären Struktur angeordnet.



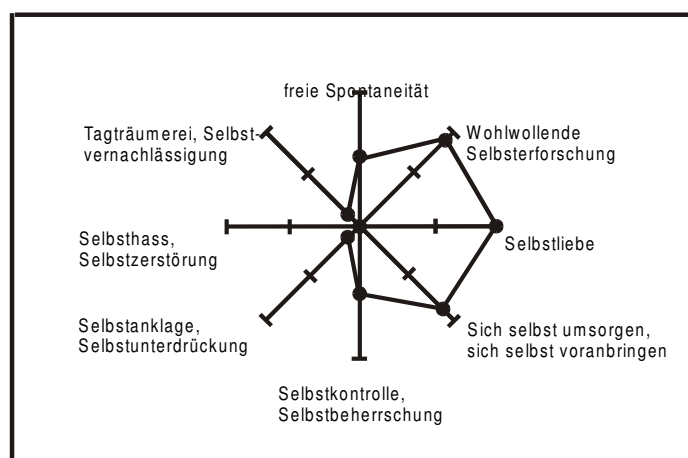
Kritik am SASB Ansatz wurde 1996 von H. Strupp, der ehemals ein Verfechter der SASB-Methode war und sie in seinen Vanderbilt-Psychotherapieforschungsprojekten eingesetzt hatte, geäußert. Er kommt im Hinblick auf das Problem geeigneter Forschungsinstrumente nun zu dem Schluss, dass selbst „so ein ausgeklügeltes System wie die Strukturelle Analyse Sozialen Verhaltens nicht die Essenz dessen erfassen kann, was wir letztlich verstehen müssten“ (Stuhr, 2001, S. 145 ).

Es gibt eine Reihe von Darstellungsmöglichkeiten der Ratings. Dies sind die Kohärenzwerte, Clustermittelwerte, die Patternkoeffizienten (ATK, CON, CFL) und die gewichteten Achsenwerte (Affiliation und Interdependenz).

### *Die Kohärenzwerte*

Die Kohärenzwerte sind ein Maß für die Stimmigkeit bzw. die interne Konsistenz eines Ratings. Die verschiedenen Qualitäten des Verhaltens lassen sich im Rahmen des SASB-Modells als eine in sich geschlossene Kurve auf einem Kreis anordnen. Der Kreis wird hierbei um den Mittelpunkt zweier Achsen gelegt: der affiliativen Achse (Freundlichkeit vs. Feindseligkeit) und der Achse der Interdependenz (Dominanz vs. Autonomie gewähren). Ähnliche Verhaltensweisen liegen nebeneinander, während gegenteiliges Verhalten um 180 Grad gegenüber auf der Kreisstruktur liegt. Zur Veranschaulichung zeigt Abbildung 28 wie der freundliche Umgang mit sich selbst durch die acht Cluster des Intrex-Fragebogens erfasst und dargestellt wird. Dieses exemplarische Rating ist in sich konsistent bzw. kohärent, da die Merkmalsausprägungen der einzelnen Cluster um so geringer sind, je weiter sie vom Cluster „Selbstliebe“ entfernt sind. Die interne Konsistenz (Kohärenz) wird durch einen Korrelationskoeffizienten ausgedrückt. Je stimmiger ein Rating ist, desto höher ist der Kohärenzwert und je widersprüchlicher ein Rating ist, desto niedriger ist er.

**Abb. 28: Beispiel eines kohärenten Ratings für einen freundlichen Umgang mit sich selbst mit kreisförmiger Anordnung der acht Cluster (SASB-Modell, intrapsychischer Fokus, aus Kammerer, 1998)**



**Tabelle 90: Mittelwerte und Standardabweichungen der Kohärenzwerte des Introjekts,  
Vergleich unserer Studie mit den Erhebungen von Kammerer (1998)  
und Davies-Osterkamp (1993)**

Kohärenz- werte	unsere Studie		Patienten *		Studenten *		Patienten A **		Patienten E **		Therapeuten **	
	M	S	M	S	M	S	M	S	M	S	M	S
Introjekt in guten Zeiten	0,88	0,19	0,74		0,91		0,50	0,61	0,61	0,51	0,86	0,39
in schlechten Zeiten	0,69	0,33	0,68		0,64		0,40	0,64	0,44	0,63	0,73	0,43

\* Davis-Osterkamp et al. in Tress 1993, S. 199

\*\* Kammerer (1998), Patienten (Anfang), Patienten (Ende)

Die Tabelle zeigt, dass die Durchschnittswerte in unserer Studie im Vergleich zu den beiden anderen Studien sowohl für das Introjekt in guten Zeiten als auch das Introjekt in schlechten Zeiten hoch sind. Der Wert für das Introjekt in guten Zeiten ist ähnlich hoch wie bei den Therapeuten in der Kammerer-Studie und den Studenten in der Düsseldorfer Studie (also den „Gesunden“) und liegt sowohl höher als der Wert für die Patienten (am Anfang und Ende der Behandlung) in der Düsseldorfer als auch in der Kammerer-Studie. Unser Durchschnittswert für das Introjekt in schlechten Zeiten ist größer als bei den Patienten zu Anfang und Ende der Behandlung bei Kammerer und ähnlich wie das der Therapeuten bei Kammerer und der Patienten und Studenten in der Düsseldorfer Studie.

Die Zahlen machen deutlich, dass unsere Patienten nach Ablauf der Behandlung eine konsistente, relativ widerspruchsfreie – „normale“ - Beziehung zu sich selbst als Folge der differenzierten Beschäftigung mit sich in der analytischen Behandlung entwickelt haben, wobei die Beziehung zu sich selbst in guten Zeiten besser ist als in schlechten Zeiten.

### *Die Clustermittelwerte*

#### **Inhaltliche Bedeutung der acht Cluster (nach Tress & Junkert, 1993)**

<b>Cluster</b>	<b>intrapsychisch (Introjekt)</b>
<b>1</b>	freie Spontaneität
<b>2</b>	wohlwollende Selbsterforschung
<b>3</b>	Selbstliebe, Selbstwertschätzung
<b>4</b>	sich selbst umsorgen, sich voranbringen
<b>5</b>	Selbstkontrolle, Selbstunterdrückung
<b>6</b>	Selbstanklage, Selbstunterdrückung
<b>7</b>	Selbsthass, Selbstzerstörung (aktiv/passiv)
<b>8</b>	Tagträumerei, desinteressierte Selbstvernachlässigung

**Tabelle 91: Clustermittelwerte und Standardabweichungen der Introjekte**

<b>Gute Zeiten</b>								
	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7	Cluster8
N	57	57	57	57	57	57	57	57
Mittelwert	67	72	72	77	42	25	8	18
Median	70	80	80	80	40	20	0	10
Standardabweichung	23	19	20	18	26	24	11	24

<b>Schlechte Zeiten</b>								
	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Cluster6	Cluster7	Cluster8
N	57	57	57	57	57	57	57	57
Mittelwert	32	48	45	54	52	46	31	38
Median	30	50	40	50	50	50	20	30
Standardabweichung	24	29	26	27	31	33	32	28

Tabelle 91 zeigt, dass sich in unserer Studie das Introjekt in guten Zeiten von dem Introjekt in schlechten Zeiten dahingehend unterscheidet, dass in guten Zeiten die Cluster 1 bis 4 deutlich höher liegen, während die Cluster 6 bis 8 deutlich niedriger sind. Cluster 5 (Selbstkontrolle) enthält in guten wie in schlechten Zeiten ungefähr gleiche Werte. Das heißt, in guten Zeiten gehen unsere Patienten erwartungsgemäß spontaner mit sich um (Cluster 1), sind wohlwollender in ihrer Selbsterforschung (Cluster 2), schätzen ihren Selbstwert höher ein (Cluster 3) und umsorgen sich selbst (Cluster 4). Der zirkumplexen Strukturen des Intrex entsprechend sind Cluster 6 (Selbstanklage), Cluster 7 (Selbsthass), Cluster 8 (Selbstvernachlässigung) in guten Zeiten nur gering ausgeprägt.

Bei dem Introjekt in schlechten Zeiten ist der Unterschied zwischen den Clustern 1 bis 4 und den Clustern 6 bis 8 weniger stark ausgeprägt.

**Tabelle 92: Clustermittelwerte (M) und Standardabweichungen (s) der Introjekte:  
Vergleich mit den Werten von Kammerer (1998), Düsseldorf  
(Davies-Osterkamp et al., 1993) und Benjamin (1988)**

	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6		Cluster 7		Cluster 8	
	M	s	M	s	M	s	M	s	M	s	M	s	M	s	M	s
<i>unsere Studie (N=56)</i>																
in guten Zeiten	67	23	72	19	72	19	77	18	42	26	25	20	8	11	18	24
in schlechten Zeiten	32	24	48	29	45	26	54	27	52	31	46	33	31	32	38	28
<i>Kammerer</i>																
<b>Pat. Anfang (N=192)</b>																
in guten Zeiten	52	28	61	28	51	28	60	25	60	28	39	31	22	26	32	28
in schlechten Zeiten	28	28	39	29	28	24	34	26	50	31	57	33	44	32	52	31
<b>Pat. Ende (N=192)</b>																
in guten Zeiten	54	29	64	26	59	25	65	23	55	27	30	27	31	23	27	25
in schlechten Zeiten	31	26	38	27	32	25	38	25	50	27	55	30	42	30	48	30
<b>Therapeuten (N=11)</b>																
in guten Zeiten	73	14	74	23	65	22	75	14	45	19	7	7	4	5	14	23
in schlechten Zeiten	30	21	33	18	31	19	43	20	57	21	65	16	43	23	47	26
<i>Düsseldorf</i>																
<b>Patienten (N=51)</b>																
in guten Zeiten	58		54		57		64		44		28		20		24	
in schlechten Zeiten	25		23		18		32		45		65		52		56	
<b>Studenten (N=101)</b>																
in guten Zeiten	56		73		70		75		51		15		4		11	
in schlechten Zeiten	33		42		38		52		51		49		27		29	
<i>Benjamin</i>																
<b>Studenten 1 (N=41)</b>																
in guten Zeiten	59	32	66	21	65	28	80	20	55	27	30	25	16	23	13	17
in schlechten Zeiten	47	33	39	22	28	24	37	23	42	24	70	26	47	34	37	30
<b>Studenten 2 (N=39)</b>																
in guten Zeiten	45	28	69	31	71	26	79	24	51	26	30	31	15	19	25	25
in schlechten Zeiten	44	28	46	29	42	25	48	27	51	30	50	31	45	33	36	55

Die Tabelle 92 zeigt unsere Studie im Vergleich zu den anderen Studien. Dabei ergibt sich folgendes Bild: In guten Zeiten haben unsere Patienten in Cluster 1 bis 4 durchweg höhere Werte als alle Vergleichsgruppen, in Cluster 6 bis 8 sind unsere Werte niedriger als die der Patienten von Kammerer zum Ende der Behandlung, sie haben ein ähnliches Niveau wie das der Düsseldorfer Patienten und der beiden Studentengruppen von Benjamin. Niedrigere Werte haben die Therapeuten der Kammerer-Studie und die Düsseldorfer Studenten.

Unser Wert in Cluster 5 ist niedriger als der aller Vergleichstichproben.

Beim Introjekt in schlechten Zeiten sind die Zahlen unserer Studie in den Clustern 1 bis 4 z. T. höher als die der Patienten zum Anfang der Behandlung und zum Ende der Behandlung bei Kammerer, sie sind auch höher als die Zahlen der Therapeuten in der Kammerer-Studie und die der Düsseldorfer Patienten.

Zu den Düsseldorfer Studenten sind die Unterschiede marginal ebenfalls zu der zweiten Studententstichprobe von Benjamin, wohingegen die erste Studententstichprobe von Benjamin eher niedrigere Werte aufweist als unsere Patienten. In Cluster 6 bis 8 beim Introjekt in schlechten Zeiten liegen die Zahlen unserer Patienten durchweg unter denen der Stichproben der Vergleichsuntersuchungen, mit Ausnahme der Düsseldorfer Studenten (Davies-Osterkamp et al. in Tress, 1993) und für Cluster 8 auch der beiden Studentengruppen von Benjamin (1988). In Cluster 5 ist unser Wert ungefähr gleich groß wie in allen anderen Studien, mit Ausnahme der ersten Studentengruppe von Benjamin, die einen niedrigeren Wert für Selbstkontrolle aufweist.

### *Die Pattern*

Um die Informationen aus den 8 Clustern eines Items des Intrex-Fragebogens zusammen zu erfassen und damit den Nachteil der Darstellung mittels der Clustermittelwerte auszugleichen, wo ein Item durch acht Werte repräsentiert wird, wurden die Patternkoeffizienten entwickelt.

#### – Der Patternkoeffizient Attack (ATK)

Der ATK-Koeffizient beschreibt, in welchem Ausmaß die Ankreuzungen des Patienten entlang der horizontalen Achse (Cluster 3 bis 7) oder der benachbarten Achse (Cluster 2 bis 6) ausgerichtet sind. Ein positiver ATK-Koeffizient bezeichnet beim Introjekt ein eher hohes Maß an Aggressivität in der Beziehung zu sich selbst, ein negativer ATK-Koeffizient bezeichnet entsprechend ein geringes Maß an Aggressivität.

#### — Der Patternkoeffizient Control (CON)

Wege der Bimodalität der CON-Koeffizienten ist eine Interpretation der Werte allein über die mittleren Werte irreführend. Um die Werte interpretieren zu können, ist die Häufigkeit positiver und negativer Control Pattern zu bestimmen. Positive CON-Koeffizienten kennzeichnen beim Introjekt einen eher kontrollierenden Umgang mit sich selbst, negative CON-Koeffizienten stehen für freie Spontaneität im Umgang mit sich selbst.

#### – Der Patternkoeffizient Konflikt (CFL)

Der Patternkoeffizient Konflikt ist wie der Patternkoeffizient CON stark bimodal verteilt, so dass auch hier die Häufigkeit positiver und negativer Pattern zu bestimmen ist. CFL plus bedeutet Konflikt zwischen Autonomie und Kontrolle/Unterwerfung, CFL minus Konflikt zwischen liebevoller Annäherung und feindseliger Ablehnung.

**Tabelle 93: Mittelwerte (M) und Standardabweichungen (s) der ATK-Koeffizienten der Introjekte verschiedener Stichproben (Kammerer, 1998; Davies-Osterkamp et al., 1993; Benjamin, 1998) im Vergleich mit unserer Studie**

ATK-Koeffizient	Introjekt in guten Zeiten		Introjekt in schlechten Zeiten	
	M	s	M	s
<b>Stichprobe</b>				
unsere Studie (N=57)	-0,78	0,23	-0,23	1,46
Patienten (A) (Kammerer) (N=192)	-0,39	0,54	0,30	0,56
Patienten (E) (Kammerer) (N=192)	-0,51	0,50	0,20	0,63
Therapeuten (Kammerer) (N=11)	-0,81	0,22	0,44	0,54
Patienten (Düsseldorf) (N=51)	-0,49		0,50	
Studenten (Düsseldorf) (N=101)	-0,82	0,19	-0,11	0,66
Studenten 1 (Benjamin) (N=41)	-0,62	0,49	-0,02	0,69
Studenten 2 (Benjamin) (N=39)	-0,65	0,32	0,24	0,55

Tabelle 93 zeigt, dass beide ATK-Werte für gute und schlechte Zeiten negativ sind, d. h. der Umgang mit sich selbst ist von Freundlichkeit bestimmt. In guten Zeiten gehen die Patienten freundlicher mit sich um als in schlechten Zeiten.

Im Vergleich zu den anderen Studien zeigt sich, dass in guten Zeiten unsere Patienten günstigere Ergebnisse, (d. h. Freundlichkeit im Umgang) zeigen als die Daten in den Vergleichstudien. Werte unserer Studie liegen nahe bei den Studenten aus Düsseldorf (1993) und den Therapeuten (Kammerer, 1998). Sie sind sowohl niedriger als die Werte der Patienten zum Anfang der Behandlung bei Kammerer als auch bei den Patienten zum Ende der Behandlung bei Kammerer.

Der Wert für das Introjekt in schlechten Zeiten ist in unserer Studie niedriger als in allen anderen Studien. Im Vergleich zu den Ergebnissen der anderen Studien gehen unsere Patienten in schlechten Zeiten freundlicher mit sich um als die Patienten bzw. die „Normalen“ in den anderen Studien.

**Tabelle 94: Mittelwerte (M) und Standardabweichungen (s) der CON-Koeffizienten der Introjekte verschiedener Stichproben (Kammerer, 1998; Davies-Osterkamp et al., 1993; Benjamin, 1988) im Vergleich mit unserer Studie**

CON-Koeffizient	Introjekt in guten Zeiten		Introjekt in schlechten Zeiten	
	M	s	M	s
<b>Stichprobe</b>				
unsere Studie	0,42	0,36	0,19	1,43
Patienten (A) (Kammerer) (N=192)	0,34	0,45	0,08	0,53
Patienten (E) (Kammerer) (N=192)	0,41	0,41	0,14	0,63
Therapeuten (Kammerer) (N=11)	0,42	0,28	0,35	0,44
Patienten (Düsseldorf) (N=51)	0,33		0,03	
Studenten (Düsseldorf) (N=101)	0,54	0,32	0,38	0,52
Studenten 1 (Benjamin) (N=41)	0,50	0,43	0,04	0,57

Tabelle 94: Positive CON-Koeffizienten kennzeichnen beim Introjekt einen eher kontrollierenden Umgang mit sich selbst, während negative CON-Koeffizienten mehr für freie Spontaneität im Umgang mit sich selbst stehen.

Beim Introjekt in guten Zeiten sind unsere Werte ungefähr gleich mit denen von Kammerers Patienten zum Ende der Behandlung und gleich den Therapeuten. Die drei Studentengruppen (Benjamin, 1988; Davies-Osterkamp et al., 1993) zeigen dagegen einen freieren und spontaneren Umgang mit sich selbst. Im Introjekt in schlechten Zeiten zeigt sich, dass unsere Patienten in schlechten Zeiten kontrollierter mit sich umgehen als in guten Zeiten. Dies ist auch in den anderen Studien festzustellen, wobei unsere Patienten relativ weniger kontrollierend mit sich umgehen als die Patienten von Kammerer (1998) zu Anfang und Ende der Behandlung und die Düsseldorfer Patientengruppe (Davies-Osterkamp et al., 1993) sowie auch die beiden Studentengruppen von Benjamin (1988). Lediglich die Therapeuten und die Studentengruppe aus Düsseldorf gehen in schlechten Zeiten freier, spontaner mit sich um als unsere Patienten.

**Tabelle 95: Mittelwerte (M) und Standardabweichungen (s)  
der CFL-Koeffizienten der Introjekte verschiedener Stichproben  
(Kammerer, 1998; Davies-Osterkamp et al., 1993; Benjamin, 1988)  
im Vergleich mit unserer Studie**

CFL-Koeffizient	Introjekt in guten Zeiten		Introjekt in schlechten Zeiten	
	M	s	M	s
Stichprobe				
unsere Studie	0,14	0,29	0,12	0,43
Patienten (A) (Kammerer) (N=192)	0,34	0,36	0,30	0,42
Patienten (E) (Kammerer) (N=192)	0,29	0,35	0,26	0,43
Therapeuten (Kammerer) (N=11)	0,19	0,28	0,35	0,27
Patienten (Düsseldorf) (N=51)	0,16		0,09	
Studenten (Düsseldorf) (N=101)	0,21	0,26	0,21	0,32
Studenten 1 (Benjamin) (N=41)	0,20	0,33	0,14	0,41
Studenten 2 (Benjamin) (N=39)	0,15	0,30	-0,02	0,36

Tabelle 95 macht deutlich, dass sich keine Unterschiede hinsichtlich des Umgangs mit sich selbst in guten Zeiten und in schlechten Zeiten zeigen. Es besteht eher ein Konflikt zwischen Autonomie und Kontrolle/Gefügigkeit (CFL positiv) als zwischen liebevoller Annäherung und feindseliger Ablehnung (CFL negativ). Im Vergleich zu den anderen Studien zeigt sich, dass der Konflikt zwischen Autonomie und Kontrolle/Gefügigkeit in unserer Studie eine geringere Rolle spielt. Beim Introjekt in schlechten Zeiten ist der positive Wert allerdings in der Düsseldorfer Patientenstichprobe (Davies-Osterkamp et al., 1993) und bei der zweiten Studentenstichprobe von Benjamin (1988) geringer als in unserer Untersuchung.

### Die gewichteten Achsenwerte

Neben den Patternkoeffizienten lassen sich die Informationen aus den Clusterwerten auch über die gewichteten Achsenwerte (Affiliation und Interdependenz) zusammenfassen.

Entsprechend dem SASB Modell lassen sich die acht Cluster auf einer Kreisfläche (zirkumplexe Struktur) anordnen. Die senkrechte Achse wird als die Interdependenz-Achse bezeichnet und die waagerechte Achse als die Achse der Affiliation. Um die Achsenwerte zu ermitteln, werden die Werte der einzelnen Cluster auf die Achse projiziert und dann addiert. Die Affiliationsachse entspricht inhaltlich dem ATK-Koeffizienten, die Interdependenz-Achse dem CON-Koeffizienten. Der Vorteil der Achsenwerte besteht darin, dass diese Werte normalverteilt sind und nicht wie die Patternkoeffizienten bimodal, sodass die Mittelwerte interpretiert werden können. Die Achsenwerte sind im Vorzeichen umgekehrt gepolt wie die Patternkoeffizienten. Positive Werte auf der Affiliations-Achse (entspricht ATK-Koeffizient negativ) bestimmen die Beziehung als eher freundlich und negative Werte (entspricht ATK-Koeffizient positiv) bestimmen sie als eher feindselig. Positive Interdependenz-Werte beschreiben eine Beziehung als eher Autonomie gewährend bzw. sich separierend (entspricht CON-Koeffizient negativ), während negative Interdependenz-Werte (entspricht CON-Koeffizient positiv) eher kontrollierendes Verhalten bzw. gefügiges Reagieren ausdrücken.

#### – Affiliation

Die Affiliationswerte liegen zwischen den Extrempolen Affiliation und Disaffiliation. Hohe Werte bedeuten im Introjekt liebevolle Freundlichkeit sich selbst gegenüber, niedrige Werte feindseliger Selbsthass.

#### – Interdependenz

Hohe Interdependenz-Werte stehen im Introjekt für ein Zulassen der eigenen Spontaneität. Niedrige Werte im Introjekt entsprechen einem Ausüben von Selbstkontrolle.

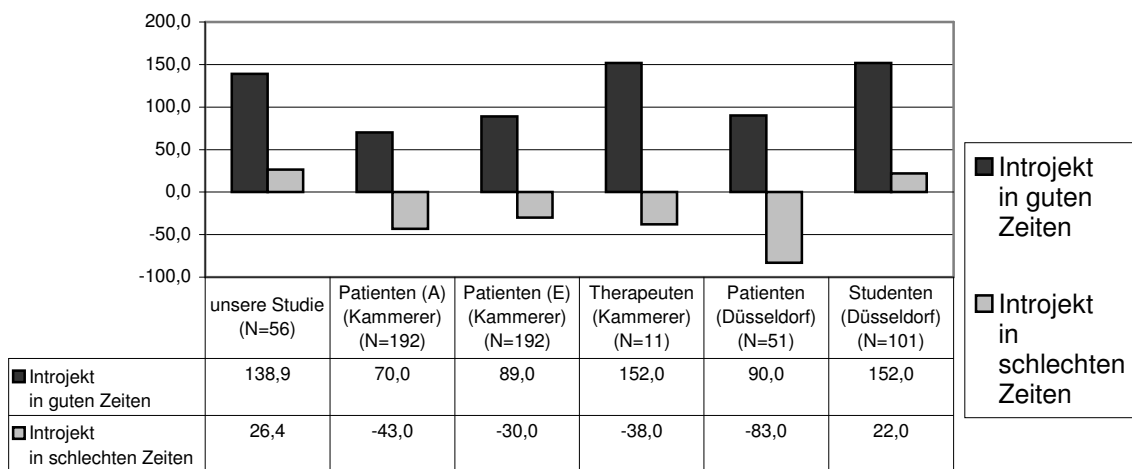
**Tabelle 96: Affiliationswerte der Introjekte  
verschiedener Stichproben  
(Kammerer, 1998; Davies-Osterkamp et al., 1993)  
im Vergleich mit unserer Studie**

Affiliationswerte	Introjekt in guten Zeiten		Introjekt in schlechten Zeiten	
	M	s	M	s
<b>Stichprobe</b>				
unsere Studie (N=56)	138,9	52,28	26,4	96,21
Patienten (A) (Kammerer) (N=192)	70	84	-43	96
Patienten (E) (Kammerer) (N=192)	89	85	-30	100
Therapeuten (Kammerer) (N=11)	152	59	-38	68
Patienten (Düsseldorf) (N=51)	90	91	-83	89
Studenten (Düsseldorf) (N=101)	152	90	22	94



Die Tabelle 96 zeigt, dass unsere Mittelwerte im positiven Bereich liegen, sowohl beim Umgang mit sich selbst in guten als auch in schlechten Zeiten. Dies bedeutet, dass die Beziehung zu sich selbst sowohl in guten als auch in schlechten Zeiten überwiegend als freundlich erlebt wird. Unsere Werte in guten Zeiten liegen deutlich über den Werten der anderen Patientenstichprobe, jedoch unter dem Wert der Therapeutenstichprobe von Kammerer (1998). Die Werte in schlechten Zeiten sind in allen anderen Stichproben im negativen Bereich, das bedeutet, dass unsere Patienten auch in schlechten Zeiten die Beziehung zu sich selbst als freundlicher erleben als die Patienten und Therapeuten in den anderen Studien.

**Abb. 29: Affiliationswerte der Introjekte unserer Studie mit verschiedenen Stichproben (Kammerer, 1998; Davies-Osterkamp et al., 1993)**



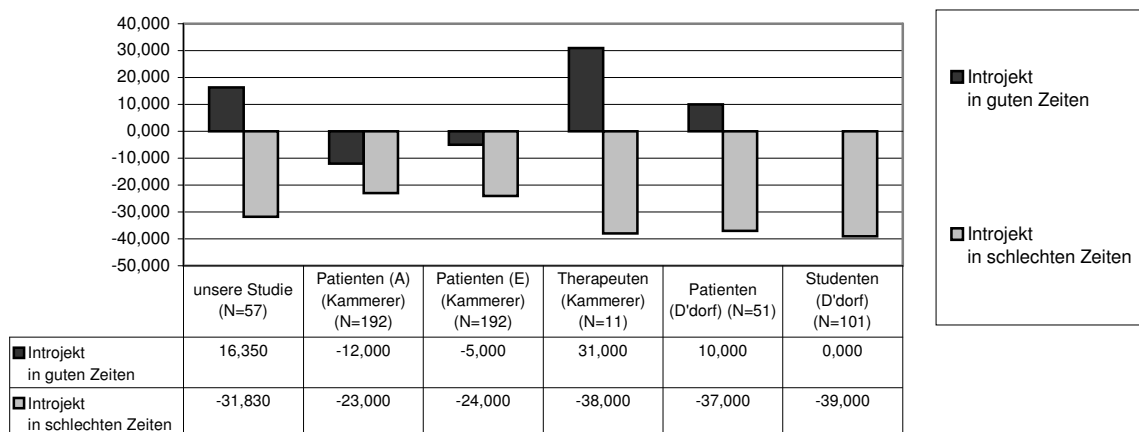
Dass unsere Patienten sich als deutlich freundlicher im Umgang mit sich selbst schildern als die Patienten der anderen Stichproben, könnte dadurch bedingt sein, dass die Patienten einer Universitätsklinik und einer psychosomatischen Rehabilitationsklinik kränker als die in ambulanten Praxen sind, unsere Patienten auch länger in Behandlung waren und damit nachhaltiger im heilsamen Kontakt zu ihren Therapeuten.

**Tabelle 97: Interdependenzwerte der Introjekte verschiedener Stichproben (Kammerer, 1998, Davies-Osterkamp et al., 1993) im Vergleich zu unserer Studie**

Interdependenzwerte	Introjekt in guten Zeiten		Introjekt in schlechten Zeiten	
	M	s	M	s
Stichprobe				
unserer Studie (N=57)	16,35	48,84	-31,83	57,50
Patienten (A) (Kammerer) (N=192)	-12	56	-23	59
Patienten (E) (Kammerer) (N=192)	-5	51	-24	51
Therapeuten (Kammerer) (N=11)	31	17	-38	68
Patienten (Düsseldorf) (N=51)	10	46	-37	66
Studenten (Düsseldorf) (N=101)	0	47	-39	56

Die Tabelle 97: Die Interdependenzwerte sind entgegengesetzt gepolt wie die CON-Koeffizienten. Hohe Interdependenzwerte stehen für ein eher Autonomie gewährendes Verhalten bzw. eine Reaktion des Sich-Separierens. Niedrige Werte entsprechen einem eher kontrollierenden Verhalten bzw. einer gefügigen Reaktion. Die Interdependenzwerte unsere Studie sind für den Umgang mit sich selbst in guten Zeiten höher als in schlechten Zeiten. Das heißt in schlechten Zeiten gehen unsere Patienten eher kontrollierend mit sich um, während sie in guten Zeiten sich selbst Autonomie gewähren. Der Vergleich mit der anderen Patientengruppe zeigt, dass unsere Patienten in guten Zeiten gewährender mit sich umgehen, sie erreichen jedoch nicht die Werte der Therapeuten (Kammerer, 1998), die das höchste Maß an Autonomie währendem Umgang mit sich selbst zeigen. In schlechten Zeiten kontrollieren sich unsere Patienten stärker als die Patienten von Kammerer, jedoch nicht so stark wie die Therapeuten und die Düseldorfer Patienten und Studenten (s. auch Abb. 30).

**Abb. 30: Interdependenzwerte der Introjekte verschiedener Stichproben (Kammerer, 1998; Davies-Osterkamp et al., 1993) im Vergleich mit unserer Studie**



*Zusammenfassend* zu den Ergebnissen im Intrex-Fragebogen kann festgestellt werden, dass unsere Patienten im Umgang mit sich selbst in guten Zeiten wohlwollender/gewährender mit sich sind als die Patientengruppen der Vergleichsstudien. Sie reichen in vielen Fällen an die „normalen“ Gruppen heran. In schlechten Zeiten gehen unsere Patienten – wie auch die Vergleichsgruppen – kontrollierter mit sich um und tun dies manchmal auch stärker als Patientengruppen von Vergleichsstudien.

Die Beziehung zu sich selbst wird in guten und in schlechten Zeiten überwiegend als freundlich erlebt. Darin schneiden unsere Patienten besser ab als die meisten Vergleichsgruppen und kommen auch hier den Gruppen der „Normalen“ (Studenten bzw. Therapeuten) nahe.

### 11.2.4 Drei Fragen zur Freude in der Therapie

(Der Fragebogen befindet sich im Anhang Kapitel 14.)

Neben Intrex-Kurzform und HAQ haben wir den Katamneseteilnehmern noch drei Fragen zur Freude in der Therapie vorgelegt, um herauszufinden, welche Rolle freudige Momente in der Therapie gespielt haben. Uns interessierte, wie häufig sich die Patienten für sich oder zusammen mit ihrem Therapeuten gefreut haben.

Zur Bedeutung der Freude in der Psychoanalyse hat Heisterkamp (1999 a u. b) eine inhaltsanalytische Untersuchung vorgelegt, in der er die stiefmütterliche Behandlung des Lust- und Freudeerlebens in Publikationen der Psychoanalyse feststellt.

Dabei spielen nach Meinung von Heisterkamp die gefühlsgeladenen und -getragenen Wirkungszusammenhänge in der analytischen Praxis eine größere Rolle als es in den gedruckten Äußerungen von Psychoanalytikern zum Ausdruck kommt.

Freudige Momente in der Therapie, vom Patienten für sich oder zusammen mit dem Therapeuten erlebt, sind nach unserer Vermutung in erfolgreich verlaufenden Behandlungen selbstverständlich anzutreffen. Sie begünstigen wahrscheinlich Behandlungserfolg bzw. -zufriedenheit und sind natürliche Begleiterscheinungen einer guten, heilsamen Beziehung zwischen Patient und Therapeut.

**Tabelle 98: Fragen zur Freude in der Therapie, Häufigkeiten und Prozente**

1. Wie häufig haben Sie sich in der Therapie gefreut?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	sehr oft	10	17,5	17,5	17,9
	hin und wieder	33	57,9	57,9	76,8
	selten	11	19,3	19,6	96,4
	sehr selten	1	1,8	1,8	98,2
	nie	1	1,8	1,8	100,0
	Gesamt	56	98,2	100,0	
Fehlend		1	1,8		
Gesamt		57	100,0		

2. Wie häufig haben Sie sich während der Therapie gemeinsam mit dem Therapeuten gefreut?

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	sehr oft	5	8,8	8,9	8,9
	hin und wieder	32	56,1	57,1	66,1
	selten	14	24,6	25,0	91,1
	sehr selten	4	7,0	7,1	98,2
	nie	1	1,8	1,8	100,0
	Gesamt	56	98,2	100,0	
Fehlend		1	1,8		
Gesamt		57	100,0		

3. Schildern Sie bitte möglichst konkret eine freudige Situation aus der Therapie.  
(Die wörtlichen Antworten befinden sich im Anhang).

Tabelle 98: Über 3/4 der Katamnesepatienten haben sich in der Therapie zumindest hin und wieder einmal gefreut, zusammen mit dem Therapeuten haben sich 2/3 der Patienten während der Therapie gefreut. Die Kategorie „hin und wieder“ wird von über 55 % der Katamnese Teilnehmer angekreuzt, sowohl was die Freude für sich als auch die gemeinsame mit dem Therapeuten angeht.

Viele Patienten erleben also im Laufe ihrer Behandlung immer mal wieder freudige Momente für sich alleine oder zusammen mit dem Therapeuten.

Im weiteren Verlauf interessieren uns einige ausgewählte Zusammenhänge zwischen den Ergebnissen in den Testinstrumenten der Verlaufsstudie und einzelnen Katamneseergebnissen einerseits und andererseits zwischen Einzelergebnissen aus der Katamnese studie.

### 11.2.5 Korrelationen

Wir untersuchen im folgenden zehn Fragestellungen (Tabellen 99 bis 108):

1) Zunächst stellten wir uns die Frage, ob es eine Beziehung gibt zwischen dem Umgang mit sich selbst (Wirkungskriterium) nach Beendigung der Behandlung und der Stärke der Beschwerden von Beginn bis Ende der Behandlung.

Die Annahme ist gerechtfertigt, dass der Beschwerdedruck (= Leidensdruck) vor, während oder am Ende der Behandlung und der Umgang mit sich selbst nach der Behandlung in unterschiedlich starkem Zusammenhang steht.

**Tabelle 99: Zusammenhang zwischen den Achsenwerten der Introjekte und den GBB-Skalen zu den Zeitpunkten t1, t3 und t5**  
(Pearson-Korrelationen)

Introjekt		GBB1			GBB2			GBB3			GBB4			GBB5		
		t1	t3	t5	t1	t3	t5	t1	t3	t5	t1	t3	t5	t1	t3	t5
		r	r	r	r	r	r	r	r	r	r	r	r	r	r	r
Affiliation	gute Zeiten	0,01	-0,29	-0,28	-0,04	0,18	-0,17	-0,10	-0,33*	-0,26	0,17	-0,16	-0,18	0,02	-0,36*	-0,32
	schlechte Zeiten	0,01	0,13	-0,02	0,21	0,17	-0,01	-0,15	-0,11	-0,19	0,20	0,05	-0,18	0,10	0,09	-0,08
Interdependenz	gute Zeiten	-0,15	-0,42**	-0,48**	-0,28	-0,36*	-0,28	-0,26	-0,51**	-0,39*	-0,10	-0,26	-0,28	-0,25	-0,57**	-0,49**
	schlechte Zeiten	-0,19	-0,12	-0,37*	-0,03	-0,14	-0,01	-0,28	-0,32*	-0,40*	-0,09	-0,16	-0,10	-0,17	-0,27	-0,32

\* Die Korrelation ist auf dem Niveau von 0,05 signifikant (2-seitig).

\*\* Die Korrelation ist auf dem Niveau von 0.01 signifikant (2-seitig).

Tabelle 99 zeigt: Der Zusammenhang mit der Stärke der Beschwerden vor dem Beginn der Behandlung (t1) und dem Umgang mit sich selbst, wie in der Katamnese geschildert, ist eher schwach ausgeprägt. Es

fällt auf, dass der Zusammenhang zwischen Interdependenz und den GBB-Skalen größer ist als zwischen Affiliation und den GBB-Skalen.

Die höchste Korrelation besteht zwischen Interdependenz in guten Zeiten und GBB-Beschwerdedruck zum Zeitpunkt t3 (- 0,57). Die Korrelation zwischen Interdependenz in guten Zeiten und Beschwerdedruck bleibt zwar auch zum Zeitpunkt t5 signifikant, der Koeffizient ist jedoch etwas niedriger als zum Zeitpunkt t3. Für die anderen Skalen des GBB gilt, dass die Korrelationskoeffizienten von t3 nach t5 eher gleich bleiben bzw. kleiner werden, mit Ausnahme Skala Erschöpfungsneigung, hier ist die Korrelation mit Interdependenz in guten Zeiten zu t5 etwas stärker als zu t3. Auch zwischen der Interdependenz in schlechten Zeiten und den GBB-Skalen lassen sich gewisse Zusammenhänge ausmachen: die Werte für Erschöpfungsneigung und Gliederschmerzen korrelieren zu t5 und t3 (Gliederschmerzen) leicht mit Interdependenz in schlechten Zeiten. Affiliation korreliert nur (allerdings schwach) in guten Zeiten mit Gliederschmerzen und Beschwerdedruck, jeweils zum Zeitpunkt t3.

Was bedeuten diese Zusammenhänge? Sich selber Autonomie gewähren in guten Zeiten nach Abschluss der therapeutischen Behandlung tritt eher dann ein, wenn die Mittelwerte zu t3 besonders in den Skalen Gliederschmerzen und Beschwerdedruck eher niedrig waren. Die Voraussage auf ein Autonomie gewährendes Introjekt nach Abschluss der Behandlung lässt sich anhand der Mittelwerte von GBB-Gliederschmerzen zu t3 sogar schon mit noch mehr Sicherheit treffen als zu t5. Auch anhand der Werte von GBB-Skala Erschöpfung zum Zeitpunkt t3 und t5 lassen sich gewisse Voraussagen auf das Autonomie-gewährende Introjekt nach Abschluss der psychotherapeutischen Behandlung treffen.

Der Zusammenhang zwischen der Freundlichkeit im Umgang mit sich selbst nach Abschluss der Behandlung und den GBB-Werten während der Behandlung ist relativ schwach. Einzig zwischen Gliederschmerzen (t3) und Beschwerdedruck (t3) und der Affiliation in guten Zeiten ergibt sich eine schwache negative Korrelation, die tendenziell besagt, dass je niedriger die Werte für Gliederschmerzen und Beschwerdedruck zum Zeitpunkt t3 sind, desto größer die Freundlichkeit im Umgang mit sich selbst in guten Zeiten nach Abschluss der Behandlung. Zwischen der Interdependenz in schlechten Zeiten und Erschöpfungsneigung, Gliederschmerzen bzw. Beschwerdedruck ergeben sich ebenfalls schwache Korrelationen, hier jeweils etwas höher zu t5 als zu t3, was vielleicht ein Hinweis darauf sein kann, dass, um ein Autonomie gewährendes Introjekt auch in schlechten Zeiten zu erreichen, eine längere Behandlung sinnvoll ist.

Zwischen den Werten der fünf GBB-Skalen zum Zeitpunkt t1 und den Achsenwerten der Introjekte ergeben sich nur geringe bis sehr geringe Korrelationen, woraus man schließen könnte, dass die anfänglichen GBB-Werte keine Voraussage zulassen auf das Introjekt nach Ende der Behandlung.

2) Ob der Grad an Entspannung im Verlauf der Behandlung in Beziehung mit dem Umgang mit sich selbst nach der Behandlung steht, untersucht die folgende Korrelation.

**Tabelle 100: Zusammenhang zwischen den Achsenwerten der Introjekte  
und dem VEV zu den Zeitpunkten t2, t3 und t5**  
(Pearson-Korrelationen)

Introjekt		VEV		
		t2 r	t3 r	t5 r
Affiliation	gute Zeiten	0,30*	0,14	0,38*
	schlechte Zeiten	0,43**	-0,03	0,35*
Interdependenz	gute Zeiten	-0,02	0,04	0,02
	schlechte Zeiten	0,22	-0,02	0,05

\* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

\*\* Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Tabelle 100 macht deutlich, dass ein gewisser Zusammenhang zwischen Entspannung, Gelassenheit, Optimismus nach halbjähriger Behandlungsdauer und der Freundlichkeit des Umgangs mit sich selbst nach Abschluss der Behandlung besteht, wobei dieser Zusammenhang noch etwas ausgeprägter ist mit der Freundlichkeit des Umgangs mit sich selbst in schlechten Zeiten ( $r = 0,43$ ). Ein geringer positiver Zusammenhang besteht auch in der Einschätzung der eigenen Entspannung und Gelassenheit nach dreijähriger Behandlungsdauer (t5) und der Freundlichkeit des Umgangs mit sich selbst sowohl in guten als auch in schlechten Zeiten ( $r = 0,38$  bzw.  $0,35$ ).

Vielleicht könnte man unter Prognoseaspekten sagen, dass der Grad an Gelassenheit und Entspannung nach halbjähriger und nach dreijähriger Behandlungszeit eine gewisse Voraussage möglich macht auf den Umgang mit sich selbst nach Abschluss der Behandlung: Je größer die Entspannung am Anfang und gegen Ende desto freundlicher der Umgang mit sich selbst nach Behandlungsende.

3) Im folgenden untersuchen wir, ob der Umgang mit sich selbst nach Abschluss der Behandlung im Zusammenhang mit den Neuroseindikatoren des MMPI(K) (Skalen Hypochondriasis, Depression, Hysterie und Psychasthenie) und ihrer Entwicklung während der Behandlung steht. Es kann angenommen werden, dass niedrige Skalenwerte (= geringer Neurotizismus) am Ende der Behandlung mit freundlicherem Umgang mit sich selbst nach Therapieende zusammengehen.

**Tabelle 101: Zusammenhang zwischen den Achsenwerten der Introjekte und den MMPI(K)-Skalen Hypochondriasis, Depression, Hysteria, Psychasthenie zu den Zeitpunkten t1, t3 und t5**  
(Pearson-Korrelationen)

Introjekt		HS			D			HY			PT		
		t1	t3	t5	t1	t3	t5	t1	t3	t5	t1	t3	t5
		r	r	r	r	r	r	r	r	r	r	r	r
Affiliation	gute Zeiten	0,14	-0,01	-0,13	-0,25	-0,24	-0,44**	0,08	-0,04	-0,20	-0,31*	-0,33*	-0,42*
	schlechte Zeiten	0,04	-0,09	-0,15	-0,18	-0,14	-0,42*	-0,03	-0,08	-0,18	-0,10	-0,08	-0,31
Interdependenz	gute Zeiten	-0,06	-0,31*	-0,30	-0,18	-0,27	-0,19	-0,06	-0,12	-0,27	-0,27	-0,35*	-0,28
	schlechte Zeiten	-0,17	-0,27	-0,14	-0,35*	-0,29	-0,15	-0,28	-0,19	-0,17	-0,24	-0,24	-0,27

\* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

\*\* Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

In Tabelle 101 wird deutlich, dass zwischen den Skalen Depression (D) und Psychasthenia (PT) des MMPI(K) ein leichter negativer Zusammenhang feststellbar ist mit der späteren Einschätzung der Freundlichkeit im Umgang mit sich selbst, insbesondere in guten Zeiten: Je geringer die Depression zum Zeitpunkt t5 und je weniger die Psychasthenia vor, nach einjähriger und nach dreijähriger Behandlung, desto freundlicher der Umgang mit sich selbst in guten und - bei D zu t5 ( $r = -0,42$ ) - in schlechten Zeiten. Ein gewisser negativer Zusammenhang besteht auch zwischen der Neigung zu Hypochondrie ( $r = -0,31$ ) nach einjähriger Behandlungszeit und der Spontaneität im Umgang mit sich selbst in guten, sowie der Stärke der Depression vor der Behandlung und der Freiheit im Umgang mit sich selbst in schlechten Zeiten ( $r = -0,35$ ). Die Skala Psychasthenia (PT) des MMPI(K) scheint am ehesten während des Therapieverlaufes geeignet, Prognosen für den Zeitraum nach Therapieende zu ermöglichen. Je geringer Zwangsverhalten im Verlauf der Behandlung ausgeprägt ist, um so freundlicher der Umgang mit sich selbst nach Abschluss der Behandlung. Dieser Zusammenhang ist unmittelbar plausibel. Von gewisser Bedeutung ist auch die Skala Depression für die längerfristige Wirkung der psychotherapeutischen Behandlung: Wer nach dreijähriger Behandlungszeit eher geringe Depressionswerte hat, kann später in schlechten Zeiten auch noch halbwegs freundlich mit sich umgehen.

4) Als nächstes untersuchen wir den Zusammenhang zwischen der Einschätzung der therapeutischen Beziehung bzw. des therapeutischen Erfolgs und des Umgangs mit sich selbst nach Beendigung der Therapie. Es ist zu vermuten, dass es hier einen starken positiven Zusammenhang gibt.

**Tabelle 102: Zusammenhang zwischen den Achsenwerten der Introjekte und den HAQ-Faktoren**  
(Spearman-Rho-Korrelationen)

Introjekt		HAQ-therapeutische Beziehung	HAQ-therapeutischer Erfolg
		r	r
Introjekt in guten Zeiten	Affiliation	0,402**	0,294*
	Interdependenz	0,295*	0,050
Introjekt in schlechten Zeiten	Affiliation	0,456**	0,376*
	Interdependenz	0,375*	0,195

\*  $p < 0,01$  \*\*  $p < 0,001$

Tabelle 102 zeigt, dass es einen positiven Zusammenhang zwischen dem Umgang mit sich selbst und der Qualität der therapeutischen Beziehung bzw. des therapeutischen Erfolgs gibt.

Für die therapeutische Beziehung sind die Zusammenhänge größer als für den therapeutischen Erfolg.

Für die Freundlichkeit im Umgang mit sich selbst (in schlechten Zeiten etwas mehr als in guten Zeiten) ist der Zusammenhang etwas ausgeprägter als für Interdependenz (Spontaneität im Umgang mit sich selbst). Für den therapeutischen Erfolg gelten die Verhältnisse entsprechend, jedoch auf niedrigerem Zusammenhangsniveau.

Das heißt, es gibt Hinweise darauf, dass ein gewisser positiver Zusammenhang besteht zwischen dem Erleben der Beziehung zum Therapeuten und der Freundlichkeit bzw. Spontaneität im Umgang mit sich selbst nach Abschluss der Behandlung. Je besser die Beziehung zum Therapeuten beurteilt wird, desto mehr Freundlichkeit und Spontaneität im Umgang mit sich selbst nach der Behandlung. Dies bestätigt Annahmen über die ursprüngliche Interaktionserfahrung mit den wichtigen Bezugspersonen korrigierende Erfahrung, die der Patient mit seinem Therapeuten gemacht hat.

5) Wir wollten auch wissen, wie sich unsere Koeffizienten im Vergleich zu denen von Kammerer (1998) gefundenen darstellen.

**Tabelle 103: Zusammenhang zwischen den Achsenwerten der Introjekte und den HAQ-Faktoren, Vergleich mit der Studie von Kammerer (1998)**

Introjekt		HAQ-therapeutische Beziehung		HAQ-therapeutischer Erfolg	
		r1	r2	r1	r2
Introjekt in guten Zeiten	Affiliation	0,402**	0,19*	0,294	0,37*
	Interdependenz	0,295*	0,20*	0,050	0,24*
Introjekt in schlechten Zeiten	Affiliation	0,456**	0,29**	0,376*	0,45**
	Interdependenz	0,375*	0,02	0,195	0,06

r1: unsere Katamnesestudie, Spearman-Rho-Korrelation

r2: Kammerer-Studie (Behandlungsende), Pearson-Korrelation

\*  $p < 0,01$

\*\*  $p < 0,001$

Tabelle 103: Die Korrelationskoeffizienten in unserer Studie zwischen therapeutischer Beziehung und Introjekt sind größer als in der Studie von Kammerer (1998). Beim therapeutischen Erfolg dagegen sind die Koeffizienten in der Kammerer-Studie größer als in unserer. Der therapeutische Erfolg steht bei Kammerer in einem größeren Zusammenhang mit dem Umgang mit sich selbst als in unserer Studie, während der Zusammenhang zwischen therapeutischer Beziehung und Umgang mit sich selbst (insbesondere Freundlichkeit in guten und schlechten Zeiten) in unserer Studie von größerer Bedeutung als bei Kammerer ist. Welche Rolle hierbei die unterschiedlichen Zeitpunkte zum Ende der Behandlung (bei Kammerer) bzw. nach Abschluss der Behandlung (unsere Studie) spielen, bleibt offen.



Vielleicht verweist dieses Ergebnis auf den besonderen Stellenwert der guten therapeutischen Beziehung in den von uns untersuchten Behandlungen für den späteren Umgang mit sich selbst.

6) Der Zusammenhang der Einschätzung von therapeutischem Erfolg und therapeutischer Beziehung im HAQ und der Grad seiner Ausprägung in unserer Studie soll im folgenden untersucht werden.

**Tabelle 104: Zusammenhang zwischen HAQ-Beziehung  
und HAQ-Erfolg in unserer Studie**  
(Spearman-Korrelation)

	HAQ-Erfolg
HAQ-Beziehung	$r = 0,673^{**}$

Tabelle 104 zeigt einen bedeutsamen, sehr signifikanten Zusammenhang zwischen dem Faktor „therapeutische Beziehung“ und „therapeutischer Erfolg“ im HAQ. Je besser die Einschätzung der therapeutischen Beziehung, um so besser die Einschätzung des therapeutischen Erfolgs und umgekehrt. Auch Bassler et al. (1995) stellten in ihren Untersuchungen mit dem HAQ die hohe Korrelation beider Skalen fest, wie ja auch der inhaltliche Zusammenhang auf der Hand liegt: Auf den empfundenen Erfolg einer Therapie nimmt die tatsächliche therapeutische Beziehung Einfluss, während umgekehrt eine gute Beziehung ohne erfolgreiches Ergebnis (bei gegebener Kompetenz des Therapeuten) nur selten auftreten dürfte. Somit bedingen sich Zufriedenheit mit dem Erfolg und dem Therapeuten gegenseitig, ohne dass dabei etwas über die Wirkrichtung ausgesagt wäre (Bassler s. o.).

An dieser Stelle kann mit Leimkühler (1996) allerdings die Frage gestellt werden, welchen Nutzen der Zufriedenheitsbegriff als Qualitätsbegriff für die Evaluationsforschung haben kann?

Alter und Geschlecht stellen z.B. Einflussfaktoren der Patientenzufriedenheit dar: ältere und weibliche Patienten(innen) sind zufriedener als jüngere und männliche.

Aus Bereichen soziologischer Zufriedenheitsforschung zeigt sich, nach Leimkühler (1996) sogar „bei Problemgruppen eine hohe Zufriedenheit bei objektiv ungünstigen Bedingungen“ (S. 767).

Sind Patienten tatsächlich zufrieden, wenn sie angeben, sie seien zufrieden ?

Es bleibe unklar, wie die Urteilsbildung von Patienten zustande kommt, ob sie sich tatsächlich auf einer Zufriedenheitsskala abbilden lasse und was die angegebene Zufriedenheit subjektiv bedeute.

Dass Zufriedenheit weniger eine Funktion objektiver Parameter als eine Funktion subjektiver Adaptation sei, zeige sich auch daran, dass generelle Verbesserungen der objektiven Lebenszufriedenheit in der Regel nicht zu höherer Zufriedenheit führen (Leimkühler, 1996, S. 771 f.).

7) Auch im Hinblick auf die Zusammenhänge zwischen den Faktoren des Helping Alliance Questionnaire (HAQ) und den fünf Skalen des GBB können wir Vergleichsstudien heranziehen. Wir möchten wissen, ob es Unterschiede zwischen unseren und den Korrelationskoeffizienten der anderen Studien gibt.

**Tabelle 105: Zusammenhang zwischen den Faktoren des Helping Alliance Questionnaire (HAQ) und den Skalen des Gießener Beschwerdeboogens (GBB)  
Vergleich mit den Ergebnissen von Kammerer (1998) und Bassler et al. (1993)**

GBB	HAQ-therapeutische Beziehung			HAQ-therapeutischer Erfolg		
	r1	r2	r3	r1	r2	r3
Erschöpfung	-0,20	-0,17		-0,12	-0,33**	
Magenbeschwerden	-0,29	-0,11		-0,09	-0,29**	
Gliederschmerzen	-0,11	-0,16		-0,09	-0,23**	
Herzbeschwerden	-0,18	-0,22*		-0,07	-0,40**	
Beschwerdedruck	-0,23	-0,20*	-0,26	-0,09	-0,37**	-0,51

r1: unsere Studie gemessen nach drei Jahren Therapie (t5); Pearson-Korrelationen

r2: Kammerer: Ende der Behandlung; Spearmann-Korrelationen, \*p < 0,01 \*\*p < 0,001

r3: Bassler et al. (1993)

Tabelle 105 zeigt, dass im Unterschied zur Studie von Kammerer (1998) keine der Skalen des GBB einen Zusammenhang mit dem Faktor „therapeutischer Erfolg“ des HAQ in der Katamnese aufweist. Mit dem Faktor „therapeutischer Erfolg“ wird das psychische Behandlungsergebnis in Form von Autonomie und Selbständigkeit erfasst, die körperliche Seite nach dreijähriger Behandlung findet dabei keine Beachtung.

Während in Bezug auf den Zusammenhang zwischen therapeutischem Erfolg und GBB unsere Ergebnisse sich von denen Kammerers (1998) und Basslers (1993) (bezogen auf den Beschwerdedruck) erheblich unterscheiden, ist die Übereinstimmung zwischen den drei Studien recht groß in Bezug auf die therapeutische Beziehung: Der Zusammenhang zwischen den Skalen des GBB und der Einschätzung der therapeutischen Beziehung ist auch in diesen Studien gering.

Die Einschätzung der therapeutischen Beziehung - und bei uns auch die Einschätzung des therapeutischen Erfolgs - nach Ablauf der Behandlung ist demnach nicht abhängig von der Ausprägung psychosomatischer Störungen nach dreijähriger Behandlungszeit. Dies könnte bedeuten, dass erlebte Symptomstärke nach immerhin dreijähriger Behandlungszeit und subjektive Zufriedenheit mit dem Therapeuten bzw. dem Therapieergebnis zwei voneinander unabhängige innere Prozesse beinhalten.

Dass das subjektive Befinden des Patienten bei Behandlungsende und sein subjektives Erleben der therapeutischen Zusammenarbeit offenbar nur einen geringeren Zusammenhang haben, bestätigen auch Rudolf und Manz (1993). Ihre Befunde verweisen demgegenüber mehr auf die prognostische Bedeutung der Therapeuteneinschätzung der therapeutischen Arbeitsbeziehung.

8) Es erscheint naheliegend anzunehmen, dass es einen positiven Zusammenhang zwischen den Faktoren des HAQ nach Ende der Behandlung und dem Grad des Optimismus und der Gelassenheit nach dreijähriger Behandlungszeit gibt.

**Tabelle 106: Zusammenhang zwischen den Faktoren des  
HAQ „therapeutische Beziehung“,  
„therapeutischer Erfolg“ und dem VEV zu den  
Zeitpunkten t2, t3 und t5  
(Spearman-Rho-Korrelation)**

HAQ	VEV		
	t2 r	t3 r	t5 r
therapeutische Beziehung	0,24	0,18	0,33*
therapeutischer Erfolg	-0,02	-0,11	0,26

\* Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Tabelle 106: Es besteht ein leichter positiver Zusammenhang zwischen der Zufriedenheit mit der therapeutischen Beziehung nach der Behandlung und dem VEV-Gesamtwert nach dreijähriger Behandlungszeit. Je mehr die Patienten sich nach dreijähriger Behandlungszeit entspannt und optimistisch erleben, um so besser wird auch die therapeutische Beziehung eingeschätzt und umgekehrt. Für die Zufriedenheit mit dem therapeutischen Erfolg besteht dieser Zusammenhang in etwas abgeschwächter Form ebenfalls.

9) Wir nehmen an, dass, wer sich nach der Therapie zufrieden mit der Beziehung zum Therapeuten oder mit dem Erfolg äußert, sich auch mehr an Freude in der Therapie erinnert.

**Tabelle 107: Zusammenhang zwischen „HAQ-Zufriedenheit mit  
der Beziehung“, „HAQ-Zufriedenheit mit dem Erfolg“ und der  
Einschätzung, wie häufig sich der/die Patient/in  
in der Therapie gefreut hat  
(Spearman-Rho-Korrelation)**

	Freude in der Therapie	
	N	r
HAQ - Beziehung	56	-0,230
HAQ - Erfolg	56	-0,359**

\*\* Korrelation ist auf dem Niveau von 0,01 signifikant (2-seitig).

\* Korrelation ist auf dem Niveau von 0,05 signifikant (2-seitig).

Anmerkung:

Negative Korrelation durch entgegengesetzte Wertigkeit der Skalenpunkte 1 bis 5 bzw. 1 bis 6 in der Freude-Skala und dem HAQ.

Tabelle 107: Der Korrelationskoeffizient ist größer und auch signifikanter zwischen der Häufigkeit der Freude in der Therapie und der Zufriedenheit mit dem Erfolg als mit der Zufriedenheit mit der Beziehung.

Freude in der Therapie scheint also etwas mehr zu tun zu haben mit dem therapeutischen Erfolg als mit der therapeutischen Beziehung.

10) Vielleicht steht die gemeinsame Freude mit dem Therapeuten während der Behandlung in einem noch engeren positiven Zusammenhang mit „Zufriedenheit in der Beziehung“ als mit „Zufriedenheit mit dem Erfolg“.

**Tabelle 108: Zusammenhang zwischen den HAQ-Faktoren „Zufriedenheit mit der Beziehung“, „Zufriedenheit mit dem Erfolg“ und „Häufigkeit der gemeinsamen Freude mit dem/der Therapeuten/in“**  
(Spearman-Rho-Korrelation)

	<b>Gemeinsame Freude mit Therapeut/in</b>	
	<b>N</b>	<b>r</b>
<b>HAQ - Beziehung</b>	<b>56</b>	<b>-0,309*</b>
<b>HAQ - Erfolg</b>	<b>56</b>	<b>-0,499**</b>

\*\* Korrelation ist auf dem Niveau von 0,01 signifikant (2-seitig).

\* Korrelation ist auf dem Niveau von 0,05 signifikant (2-seitig).

Anmerkung:

Negative Korrelation durch entgegengesetzte Wertigkeit der Skalenpunkte 1 bis 5 bzw. 1 bis 6 in der Freude-Skala und dem HAQ.

Der Korrelationskoeffizient von Freude, diesmal gemeinsamer Freude mit dem Therapeuten und Zufriedenheit mit dem Erfolg ist, anders als angenommen, größer als mit der Zufriedenheit mit der Beziehung (Tabelle 108). Der Zusammenhang zwischen der Zufriedenheit mit dem Erfolg und der gemeinsamen Freude mit dem Therapeuten ist deutlich, je größer die Zufriedenheit mit dem Erfolg, desto häufiger die gemeinsame Freude mit dem Therapeuten und umgekehrt. Dieser Zusammenhang gilt nur in abgeschwächter Form auch für die Zufriedenheit mit der Beziehung. Beide Korrelationskoeffizienten sind größer als zwischen Freude allgemein und den HAQ-Faktoren, was auf die Bedeutung des freudigen Miteinanders mit dem Therapeuten insbesondere für den Erfolg der Behandlung zu verweisen scheint.

Zusammenfassend zu den Korrelationen lässt sich folgendes sagen:

- Der Störungsgrad in den körperbezogenen Beschwerden vor der Behandlung ist kein Indikator für die Qualität des Introjekts nach Abschluss der Behandlung.
- Ein geringer Beschwerdedruck und Erschöpfungsgrad, sowie größere Entspannung am Anfang und Ende der Behandlung, stehen in positivem Zusammenhang mit einem Autonomie gewährenden Introjekt und einem freundlichen Umgang mit sich selbst nach Ablauf der Behandlung.
- Geringes Zwangsverhalten und niedrige Depression nach drei Jahren Behandlung lassen nach Ende der Behandlung einen freundlichen Umgang mit sich selbst erwarten.

- Bei einem hohen Grad an Entspannung nach dreijähriger Behandlungszeit kann mit einer hohen Zufriedenheit mit der therapeutischen Beziehung nach der Behandlung gerechnet werden.
- Die erlebte Stärke der Körpersymptome zum Behandlungsende steht nicht im Zusammenhang mit subjektiver Zufriedenheit mit dem Therapeuten und dem therapeutischen Erfolg nach der Behandlung.
- Zwischen dem Erleben der Beziehung zum Therapeuten und der Freundlichkeit und dem Umgang mit sich selbst besteht ein positiver Zusammenhang, ebenso zwischen der Einschätzung der therapeutischen Beziehung und des therapeutischen Erfolges.
- Häufigkeit der Freude in der Therapie, mehr noch der gemeinsamen Freude mit dem Therapeuten, steht in einem positiven Zusammenhang mit der Zufriedenheit mit dem Erfolg.

### 11.2.6 Zusammenfassung und Kommentierung der Katamneseergebnisse

57 Patienten haben an der Katamnese teilgenommen. Bei der Hälfte liegt die Therapie, vom Katamnesezeitpunkt 2/2000 aus gerechnet, zwei bis sechs Jahre zurück, bei der anderen Hälfte bis zu zwei Jahre. 55 % der Patienten haben die Therapie in den Jahren 1997 bis 1998 beendet.

Die Geschlechterverteilung zum Zeitpunkt der Katamneseerhebung ist wie in der Untersuchungsgruppe zum Zeitpunkt t1; über 90 % sind im Alter von 31 bis 50 Jahre.

Über 3/4 stammen aus der Untersuchungsgruppe, 1/4 aus der Kontrollgruppe.

Die Teilnehmer an der Katamnese sind tendenziell älter als die Nichtteilnehmer, Männer und Frauen sind in beiden Gruppen gleich verteilt.

Die Behandlung war bei Teilnehmern und Nichtteilnehmern an der Katamnese gleich effektiv (Effektivitätsmaß t1:t5).

Zwei Verfahren kamen zum Einsatz, um die langfristige Wirkung der Behandlung zu untersuchen:

- Intrex-Fragebogen (Kurzform) von Benjamin und Tress
- Helping Alliance Questionnaire von Luborsky, deutsch von Bassler et al. (1995a)

Zusätzlich wurden den Katamneseteilnehmern drei Fragen zur Freude in der Behandlungssituation gestellt.

#### *Die Ergebnisse*

Unsere Patienten haben nach Ablauf der Behandlung, gemessen an den Werten von Vergleichsstudien, eine konsistente, relativ widerspruchsfreie – „normale“ - Beziehung zu sich selbst entwickelt.

Sie gehen so freundlich (und freundlicher) sowohl in guten als auch in schlechten Zeiten mit sich um wie die „Normalen“ und die Patienten anderer Studien am Ende ihrer Behandlung.

Im freien und spontanen Umgang mit sich selbst in guten und in schlechten Zeiten sind unsere Patienten mit den Patienten der anderen Studien gleichauf, in schlechten Zeiten sogar noch freier.

Der Konflikt zwischen Autonomie und Kontrolle/Gefügigkeit spielt in unserer Studie eine geringere Rolle als in den anderen Studien.

Die Beziehung zu sich selbst wird sowohl in guten als auch in schlechten Zeiten überwiegend als freundlich erlebt. Unsere Patienten schildern sich als deutlich freundlicher im Umgang mit sich selbst in guten und schlechten Zeiten als die Patienten der anderen Stichproben.

In schlechten Zeiten gehen unsere Patienten eher kontrollierend mit sich um, während sie in guten Zeiten sich selbst Autonomie gewähren. Der Vergleich mit der anderen Patientengruppe zeigt, dass unsere Patienten in guten Zeiten gewährender mit sich umgehen, während sie sich in schlechten Zeiten stärker kontrollieren.

Ohne die Werte vor der Behandlung zu kennen, wird man doch annehmen können, dass dieses Ergebnis der vorangegangenen psychotherapeutischen Behandlung zuzuschreiben ist. Es zeigt vor allen Dingen in guten Zeiten den gelasseneren Umgang mit sich selbst, der zu mehr Freundlichkeit und Spontaneität führt und der behandlungsbedingten Versöhnung mit den frühen Objekten der Kindheit zu verdanken sein dürfte.

Im Helping Alliance Questionnaire erhält der Satz: „Ich habe einige neue Einsichten gewonnen“ mit Abstand die größte Zustimmung.

Die Zustimmung ist am schwächsten für die Aussage „Ich sehe, dass ich die Probleme bewältigen kann, wegen derer ich zur Behandlung kam“.

Einsichtsgewinnung als therapeutischer Nutzen wird demnach von den meisten Patienten als wesentlich bedeutsamer eingeschätzt als Problembewältigung.

Unsere Patienten sind mit der Behandlung sehr zufrieden, sowohl was die therapeutische Beziehung als auch den therapeutischen Erfolg angeht. Sie schätzen die Qualität der therapeutischen Beziehung genauso hoch ein wie die Patienten der Vergleichsstudien, während der therapeutische Erfolg in unserer Studie höher eingeschätzt wird. Bei den zum Vergleich herangezogenen Studien handelt es sich um kürzere stationäre Behandlungen, denen gegenüber die Bedingungen für Veränderung in den von uns evaluierten Behandlungen in vielerlei Hinsicht günstiger sind.

Zur Freude in der Therapie: Zumindest hin und wieder gefreut haben sich allein oder zusammen mit ihrem Therapeuten über 75 % der Katamnesepatienten. Dies ist wiederum erfreulich, weil es zeigt, dass es in vielen Behandlungen auch leichte und entspannte Momente gibt.

Zwischen den Wirksamkeitsmaßen der Verlaufsstudie und den Katamneseergebnissen bzw. zwischen den Katamneseergebnissen selbst zeigen sich folgende Zusammenhänge:

- Die Faktoren „therapeutische Beziehung“ und „therapeutischer Erfolg“ im HAQ. stehen in einem bedeutsamen Zusammenhang: Je besser die Einschätzung der therapeutischen Beziehung, um so besser ist auch die Einschätzung des therapeutischen Erfolgs und umgekehrt. Man kann demzufolge annehmen, dass i. a. eine schlechte Beziehung letztlich keinen therapeutischen Erfolg hervorbringt und ein therapeutischer Misserfolg nicht mit einer guten Beziehung zum Therapeuten einhergeht.

- Die Einschätzung der therapeutischen Beziehung und des therapeutischen Erfolgs nach Ablauf der Behandlung steht in keinem Zusammenhang mit den psychosomatischen Störungen nach dreijähriger Behandlungszeit. Die Zufriedenheit mit der psychotherapeutischen Behandlung ist nicht abhängig von der Symptombelastung durch körperliche Beschwerden nach dreijähriger Behandlungszeit. Offensichtlich gibt es für Patienten andere Kriterien für die Behandlungszufriedenheit als die Symptombelastung, wie z. B. die Einsichtsgewinnung.
- Entspannung, Gelassenheit, Optimismus nach halbjähriger und nach dreijähriger Behandlungsdauer stehen in einem gewissen positiven Zusammenhang mit der Freundlichkeit des Umgangs mit sich selbst nach Abschluss der Behandlung. Vom Grad an Gelassenheit und Entspannung im Anfangsstadium und nach fortgeschrittener Langzeitbehandlung ist eine gewisse Voraussage möglich auf den Umgang mit sich selbst in der Zeit nach Abschluss der Behandlung.  
Wer schon am Anfang der Langzeitbehandlung und erst recht an ihrem Ende optimistisch ist, bleibt es auch nach Abschluss, was sich im freundlichen Umgang mit sich selbst zeigt.
- Die Skala Psychasthenie des MMPI(K) kann über den gesamten Beobachtungszeitraum als ein gewisser Gradmesser für den die Behandlung überdauernden Therapieerfolg, ausgedrückt in der Freundlichkeit im Umgang mit sich selbst, gelten: Geringe PT-Werte machen einen freundlicheren Umgang mit sich selbst nach der Behandlung wahrscheinlich. Auch die Skala Depression des MMPI(K) mit ihrem Mittelwert am Ende des dreijährigen Beobachtungszeitraumes lässt einen Zusammenhang zum Introjekt in der Katamnese erkennen: Je geringer die Depression nach dreijähriger Behandlung, desto freundlicher der Umgang mit sich selbst in schlechten Zeiten. Zwangsgedanken und Depression gehen einher mit einem wenig freundlichen Umgang mit sich selbst. Wer Zwänge und Depression im Laufe der Behandlung mildern konnte, kann nach der Behandlung freundlicher mit sich umgehen, gerade auch in schlechten Zeiten.
- Freundlichkeit im Umgang mit sich selbst in guten und schlechten Zeiten und die Einschätzung der therapeutischen Beziehung (weniger des therapeutischen Erfolges) stehen in einem gewissen positiven Zusammenhang.
- Tendenziell zeigt sich, dass je mehr die Patienten sich nach dreijähriger Behandlungszeit entspannt und optimistisch erleben, um so besser therapeutische Beziehung und - mit etwas mehr Einschränkungen - auch therapeutischer Erfolg eingeschätzt werden.
- Häufigkeit der Freude in der Therapie und Zufriedenheit mit dem therapeutischen Erfolg und der therapeutischen Beziehung stehen in einen positiven Zusammenhang. In Bezug auf den therapeutischen Erfolg ist der Zusammenhang noch deutlicher.
- Die gemeinsame Freude mit dem Therapeuten scheint insbesondere für den Therapieerfolg wichtiger zu sein als die Freude allein. Wir nehmen an, dass eine positive Entwicklung im Laufe der The-

rapie einhergeht mit mehr Freude für sich selbst oder zusammen mit dem Therapeuten, größerer Entspannung und Gelassenheit, einer besseren Beziehung zum Therapeuten, einer günstigeren Einschätzung des Therapieerfolges und schließlich einem wohl gesonnenen Introjekt, das auch in schwierigeren Zeiten Bestand hat.



## **12. Abschließende Stellungnahme zur Gesamtuntersuchung und Stellenwert im Kontext der psychoanalytischen Wirkungsforschung**

Unsere Untersuchung ist eine empirische und im ersten Behandlungsjahr auch eine vergleichende Studie. Sie ist sowohl prospektiv wie auch naturalistisch angelegt, indem sie die realen Gegebenheiten psychotherapeutischer Praxen berücksichtigt, in denen die Patienten vom Therapeuten gemäß seiner persönlichen Einschätzung und seiner Stundenkapazität entweder in Behandlung genommen, abgewiesen oder auf einen Warteplatz verwiesen werden. Dieses gängige Verfahren wurde von uns für die Bildung einer Untersuchungs- und Kontrollgruppe aufgegriffen und genutzt. Wir betrachten dieses Vorgehen als praxisorientiert und die Patientenauswahl als zufällig, auch wenn die Kriterien einer streng randomisierten Auswahl nicht voll eingehalten wurden.

Unser Kontrollgruppendesign mit einer auf ein Jahr begrenzten Vergleichsdauer von Untersuchungs- und Kontrollgruppe halten wir für vertretbar (s. auch Stuhr, 2001 b), weil es a) die realen Gegebenheiten der psychotherapeutischen Praxis lediglich aufgreift, nicht jedoch aus Forschungszwecken verändert, und b) weil wir so doch zumindest einen interessanten Zeitabschnitt einer quasi-experimentellen Kontrolle unterwerfen können. Es handelt sich hierbei um einen Kompromiss zwischen Wünschbarem, Machbarem und Vertretbarem. Zwar sind damit die vermeintlich erforderlichen Kriterien eines idealtypischen Experimental-Designs mit durchgehender Kontrollgruppe bis zum Abschluss des Quasi-Experiments, so wie dies von manchen Kritikern für nötig gehalten wird (vgl. Grawe, 1994, Fisher u. Greenberg, 1996), nicht eingehalten worden, dennoch erbringt unser Vorgehen für den beobachteten begrenzten Zeitabschnitt immerhin nachprüfbare Daten und ermöglicht nach methodologischen Gesichtspunkten abgeleitete empirische Aussagen über einen Vergleichszeitraum von einem Jahr. Und mit einigem Plausibilitätsrecht kann man doch zumindest die begründete Annahme aussprechen, dass sich der hierbei zeigende Trend auch bei einer Verlängerung des Vergleichszeitraums bestätigt hätte.

Da der Kontrollgruppenzeitraum auf ein Jahr begrenzt war, könnte man fälschlicherweise zu der Annahme kommen, aus diesen Daten Aussagen zur Wirksamkeit analytischer Kurzzeithherapie machen zu können. Dies ist jedoch nicht zulässig, da davon auszugehen ist, dass die innere Verfassung eines Kurzzeitpatienten und eines Langzeitpatienten bei Aufnahme der Therapie eine ganz andere ist, wodurch sich dann auch die Therapien an beiden Personengruppen voneinander unterscheiden dürften. Oder positivistisch ausgedrückt: die unabhängigen Variablen sind nicht gleich, weswegen die Unterschiede in der abhängigen Variable wissenschaftlich auch nicht zu interpretieren sind.

Den Patienten konnte aber nach Meinung der beteiligten Kollegen, und dies deckt sich mit der in der Literatur einhellig vertretenen Ansicht, keine längere Wartezeiten zugemutet werden, und zwar sowohl aus ethischen wie auch aus pragmatischen Gründen. Die pragmatischen Gründe bestehen darin, dass eine Kontrollgruppe sich schon deshalb nicht länger als ein Jahr aufrecht erhalten lässt, weil die betroffenen Patienten sich bei noch längeren Wartezeiten wahrscheinlich nach anderen Therapeuten umgesehen hätten.

Unser naturalistischer Ansatz ist nach dem aktuellen Stand der Wirkungsforschung zeitgemäß. Dagegen wird unser Kontrollgruppendesign auf die Ablehnung anderer Therapieforscher stoßen, so wie z.B. bei Stuhr (2001b), der dieses Design nur noch bei der Einführung neuer Verfahren für angebracht hält.

Der naturalistische Ansatz vermeidet die praxisfernen Nachteile experimenteller Wirkungsforschung (z.B. randomisierte Zuweisung zu Kontroll- oder Untersuchungsgruppe), muss dafür allerdings in Kauf nehmen, dass sich die subjektiven Kriterien des Therapeuten bei der Zuweisung zu einer der beiden Gruppen auf die Ergebnisse in den abhängigen Variablen auswirken. Dass sich unsere UG und KG zum Zeitpunkt der probatorischen Sitzungen hinsichtlich wichtiger Kriterien (Alter, Geschlecht, Diagnose) kaum unterscheiden (vgl. Kapitel 10.3.2.1), hat sich zwar rein zufällig ergeben, könnte aber ein Hinweis darauf sein, dass die behandelnden Analytiker ihre Patienten nicht nach einem subjektiven Bias in die eine oder andere Gruppe zugeteilt haben. Allerdings ist damit noch nicht gesagt, dass sich beide Gruppen nicht doch in einem ungeprüften und gleichzeitig ergebnisrelevanten Kriterium unterscheiden.

Unser Studiendesign ist an der Methodik der empirischen Sozialforschung angelehnt und ließe sich daher natürlich auch zur Überprüfung anderer Therapieverfahren anwenden. Das Studiendesign ist nicht auf den Gegenstand „Psychoanalyse“ oder „analytische Psychotherapie“ zugeschnitten und erfasst auch nicht die Besonderheiten dieser Verfahren. Dies war auch nicht die Zielsetzung unserer Studie, sondern sie sollte einen Wirksamkeitsnachweis individualpsychologisch-psychoanalytischer Einzelpsychotherapie erbringen, der einer positivistischen Kritik ‚à la Grawe‘ (1992) standhält. Insofern befindet sich unser Studiendesign unter dem Gesichtspunkt der Zuordnung zur Geschichte der Psychotherapieforschung gesehen wieder in der Phase eins, der sog. Rechtfertigungsforschung, was uns jedoch nicht stört, denn uns ging es ja darum, den dem universitären Wissenschaftsdenken bisher schuldig gebliebenen Wirksamkeitsnachweis nachzuliefern.

Was die Messverfahren angeht, die heute als geeignet angesehen werden, die Wirksamkeit von Behandlung zu überprüfen, so sind eine Reihe von Verfahren im Einsatz, mit denen patienten- und therapeuten- seitige Daten erhoben werden. Beispielhaft sei die in zahlreichen Studien (z.B. Sandell, 1999, 2000) verwendete Symptom Check Liste SCL-90-R (Derogatis et al. 1973, Schauenburg et al. 1998) erwähnt. Dies ist ein Verfahren zur Selbstbeurteilung von Patienten hinsichtlich verschiedener belastender Symptome. Der Fragebogen umfasst 90 Items, davon werden 83 zu den folgenden Symptombereichen zusammengefasst: Somatisierung, Zwanghaftigkeit, Depressivität, Ängstlichkeit, Unsicherheit im sozialen Kontakt, Aggressivität, Feindseligkeit, Phobische Angst, Paranoides Denken und Psychotizismus. Die Items werden zu einem Gesamtwert GSI (Global Severity Index) zusammengefasst. Er gibt Auskunft über das allgemeine Ausmaß psychischer Beschwerden. Ein Vorteil ist die klinische Bedeutsamkeit durch Nennung diagnostisch klar abgegrenzter Symptombereiche, die auch den Faktor Somatisierung beinhalten. Wir haben uns nicht auf dieses Verfahren gestützt, - hauptsächlich auch darum nicht, weil uns dieses Verfahren zu Beginn unserer Untersuchung noch nicht bekannt war. Die Gründe für unsere Wahlentscheidung zugunsten des MMPI(K), GBB und VEV sind oben diskutiert worden.

Unter dem Blickwinkel heutiger Forschungsstandards, insbesondere was die Auswahl der Messinstrumente angeht, kann gefragt werden, ob mit unseren Instrumenten punktuell und prozessorientiert die Wirksamkeit analytischer Psychotherapie tatsächlich gemessen werden kann. Diese Frage ist insbesondere in Bezug auf die Verwendung der MMPI-Kurzform angebracht, dem zu Forschungszwecken im klinischen Bereich, wegen der geringen Trennschärfe seiner Skalen nur noch eine eingeschränkte (bestimmte) Bedeutung beigemessen wird. Strupp (1978) beurteilte Persönlichkeitsfragebögen wie den MMPI, den er anfänglich (Anfang der 70er Jahre) auch in der sog. Vanderbilt I-Studie einsetzte, lediglich als für die globale Messung von durch Psychotherapie bewirkte Veränderungen

geeignet. Und als solchen haben wir den MMPI(K) zumeist auch genutzt, insbesondere in der Fokussierung auf die neurotische Trias+PT.

Mit dem Intrex und dem HAQ haben wir in der Katamnese qualitative Verfahren, auf analytischem Konzept aufbauend oder ihm nahestehend, eingesetzt. Dabei ist auch der Patient selbst in der Beurteilung seiner Behandlung zu Wort gekommen (HAQ). Außerdem wurden den Patienten auch Fragen zum Vorkommen von „Freude“ in der Therapie gestellt.

Für Wirksamkeitsuntersuchungen wird heute „Multidimensionalität“ gefordert. Gemeint ist damit die Ergänzung objektiver durch subjektive Daten (Selbstbeurteilung der Therapiefortschritte, Fremdbeurteilung z.B. der Therapeuten). Es werden empirische, naturalistische und prospektive Herangehensweisen in der Hoffnung miteinander gemischt, dadurch dem Gegenstand der Psychoanalyse bzw. der analytischen Psychotherapie angemessenere Designs zu finden. Sie sollen der Psychoanalyse „als Wissenschaft zwischen den Wissenschaften“ gerechter werden. So propagiert Leuzinger-Bohleber die Einzelfallstudie, in der die „Idiosynkrasie des Einzelfalles“ den ihr gemäßen Stellenwert erhält, aber in aggregierter Form auch Verallgemeinerungen zulässt. Es kann diskutiert werden, in wie weit es sich hierbei um einen wirklich neuen Ansatz oder eben doch lediglich um einen eklektizistischen Mischansatz handelt. Es fehlt nicht an Designansätzen, - das Problem ist die eigene Wissenschaftstheorie der Psychoanalyse, die der Boden für eine eigene Wissenschaftslogik darstellen könnte.

Unser Ansatz bleibt hinter der Forderung nach Multidimensionalität deutlich zurück. Unsere Therapeuten sind überhaupt nicht befragt worden, und unsere Patienten wurden zunächst lediglich quantitativen Verfahren unterzogen, - erst in der Katamnese wurden sie um die Beurteilung ihrer Behandlung gebeten. Gleichwohl sehen wir in unserem Ansatz den Vorteil, weg von allen „Einschätzungen“ und stattdessen unter Anwendung gleich gebliebener Messinstrumente zu Messwerten gekommen zu sein, deren Veränderung unter Berücksichtigung wissenschaftslogischer Ablaufsregeln nur auf die intermittierende Variable „Therapieeinfluss“ zurückzuführen ist.

Wir haben aus forschungsökonomischen Gründen nach dreijähriger Beobachtungs- und Behandlungszeit einen „Schnitt“ machen müssen. Hätten wir der Forderung nach „Multidimensionalität“ folgen wollen, hätten wir über unsere Kollegen bei den Patienten nachexplorieren und weitere Forschungen anstrengen müssen. Aus ökonomischen Gründen haben wir uns jedoch entschlossen, den Datensammlungsprozess an dieser Stelle zu beenden.

Was unsere Katamnese angeht, so liefert sie einige aufschlussreiche Ergebnisse, insbes. zum posttherapeutischen Introjekt, das sich durch die Bearbeitung hauptsächlich der Übertragungsbeziehung während der analytischen Therapie in Richtung Selbstakzeptanz verändert hat und damit die Voraussetzung für ein gedeihlicheres Leben entstanden ist.

Die positive Beurteilung der Behandlung hinsichtlich Erfolg und therapeutische Beziehung – auch im Vergleich zu anderen Studien – lässt eine günstige Prognose zu, wenn auch nicht alle Anforderungen an eine katamnestiche Studie (s. Kap. 13) von uns erfüllt wurden. Im Rahmen unseres Forschungsprojektes hätte sich ein an den oben angeführten Standards orientiertes Vorgehen nicht realisieren lassen.

Bei jeder Studie bleiben wichtige Fragen unbeantwortet. Dies gilt auch für unsere Studie. So haben wir z.B. nicht zwischen abgeschlossenen, abgebrochenen, Nonrespondern (fehlende Daten wg. Non-Compliance oder anderer Gründe) unterschieden. Hier bietet sich ein weites Feld für weitere Forschung, sowohl quantitativer als auch qualitativer Art: Welche prozessualen Bedingungen führen zu Non-Compliance, zu Abbrüchen und zu einer erfolgreichen Beendigung der Behandlung, welche personalen Bedingungen auf Seiten von Therapeut und/oder Patient sind ausschlaggebend? Gewisse Hinweise lassen sich aus unserer Untersuchung entnehmen z.B. zur unterschiedlichen Effektivität der Patient/Therapeut-Dyade (s. Kap.10.4).

Weiterhin könnte man als Maßstab für die Relevanz unserer Untersuchung und ihrer Ergebnisse Hartkamps offene Fragen im Bereich der psychoanalytischen Wirkungsforschung anlegen (s. Kap.13). Dann wird man zu dem Schluss kommen müssen, dass wir zur Beantwortung dieser Fragen wenig Aufschluss geben konnten. Wie überhaupt die Leitfrage der Phase drei („Was geschieht - mikroanalytisch betrachtet - wirklich im Behandlungsprozess?“) der Forschungsgeschichte, in der sich die Forschung heute eigentlich befinden sollte und ohne den neuen durch Grawe ausgelösten Druck sich mit eigentlich schon, jedoch offensichtlich zu früh, abgehakten Problemen der Wirksamkeit zu befassen, auch möglicherweise tatsächlich befände, von unserer Studie nur peripher berührt wird. Von der Beantwortung der Frage nach dem, was wirklich im Behandlungsprozess geschieht, was wirklich wirksam ist, ist die Psychoanalyse noch weit entfernt. Heisterkamp (2001) hat bei seinen Forschungen sog. enactments zwischen Analytiker und Patient analysiert und entdeckt dabei die Bedeutung des unmittelbaren im Gegensatz zum mittelbaren Verstehen, „dessen Wesen und Wirkung gerade in der Gegenwärtigkeit und der Unmittelbarkeit der Begegnung“ (S. 188) zwischen Analytiker und Patient liegt.

Es bleibt abzuwarten, ob und wie sich die empirische Forschung der selbstgestellten Aufgabe stellt, herauszufinden, was wirklich, mikroanalytisch betrachtet, im Behandlungsprozess geschieht.

### **13. Aspekte zukünftiger Psychotherapieforschung**

Die Psychotherapieforschung als Ergebnisforschung wird sich nach der Auffassung von Kordy & Kächele (1996) in Zukunft hauptsächlich mit Grundlagenforschung und angewandter Psychotherapieforschung beschäftigen. Dabei wird es eine Vielzahl von unterschiedlichen Forschungsansätzen geben.

In der Grundlagenforschung geht es einmal um die Erfassung des „wirklich Wirksamen“, die differenzierte Beschreibung des psychotherapeutischen Prozesses unter weitest gehendem Verzicht auf abstrahierende Reduktion. Ein Beispiel dafür gibt Leuzinger-Bohleber (1994). Die Feinanalyse psychotherapeutischer Prozesse soll entscheidend zu einem besseren Verständnis, wie Psychotherapie wirkt, beitragen.

Huber u. Kächele (2000) haben einen Schulenunterschied in Bezug auf den die outcome-Varianz aufklärenden Anteil spezifischer vs. unspezifischer Faktoren festgestellt: die Wirkung von unspezifischen Faktoren auf das Therapieergebnis wird von kognitiv-behavioral orientierten Forschern für eher gering gehalten oder in Frage gestellt.

Die Grundlagenforschung wird versuchen, die von Bozock u. Bühler (1988) gestellte Frage zu klären, ob und in welchem Maße die von den Therapieschulen entwickelten Methoden, die spezifischen Faktoren, tatsächlich darauf reduzierbar sind, dass sie lediglich jeder einzelnen Therapiesitzung aufs Neue die

Struktur geben, in der die eigentlich effektiven unspezifischen Faktoren in der Therapie zum Wirken kommen können.

Zu den unspezifischen Faktoren gehören die therapeutische Situation und die zwischenmenschliche Beziehung des Patienten zum Therapeuten, die hilfreiche Allianz als Funktion der persönlichen Eigenschaften der Therapeuten und / oder der Patienteneigenschaften, wie Ich-Stärke, Organisationsstruktur, Reife, Motivation und die Fähigkeit zum produktiven Sich-Einlassen.

Weiterhin steht in der Grundlagenforschung die Frage der Wirksamkeit für homogene Krankheitsgruppen - die geeignete Behandlung für bestimmte Gruppen von Patienten bzw. Krankheiten - im Vordergrund.

Neben der Grundlagenforschung wird die angewandte Psychotherapieforschung eine bedeutende Rolle spielen. Hier werden Fragen danach, wie viel Therapie für welche Patienten ausreichend ist, um die gewünschten Ergebnisse zu erzielen und wie man begrenzte therapeutische Ressourcen (z. B. die Sitzungen über die Behandlungsdauer) möglichst optimal verteilt zum Forschungsgegenstand. Nach Art der Dosis-Wirkungs-Modelle (Howard et al. 1986) wird die Beziehung zwischen Gesamtstundenzahl in Relation zur Dauer der Behandlung und den Ergebnissen der angewendeten Therapie für Diagnose- und Störungsgrad-Subgruppen untersucht.

Des weiteren sollen z.B. im Rahmen von Qualitätssicherung (QS) Standards für Weiterentwicklungen von Modifikationen in der Behandlungstechnik oder der Erweiterung von Indikationsbereichen gefunden werden (Kordy & Kächele, 1996).

Beziehen sich die o.g. Forschungsgegenstände auf den gesamten Bereich der wissenschaftlich anerkannten Psychotherapie, so sieht Hartkamp (1997, S.123 f.) in Bezug auf die psychoanalytische Therapie, deren Vertreter in den letzten Jahren verstärkte empirisch orientierte Forschungsbemühungen unternehmen (Huber u.Kächele, 2000) folgende noch offene Fragen:

- 1) Die Effektivität und die zeitlichen Charakteristika des Veränderungsprozesses im direkten Vergleich mit anderen psychodynamisch orientierten und verhaltenstherapeutischen Therapieansätzen in Fällen von schwerer psychischer Krankheit, bei denen langfristige psychoanalytische Behandlungen den kurz- und mittelfristigen Psychotherapien, so die Annahme, überlegen sind. In dem Zusammenhang von Effektivität spielt auch die sogenannte „strukturelle Veränderung“ als Wirksamkeitsindikator psychoanalytischer Psychotherapien eine wichtige Rolle.
- 2) Der Einfluss des psychoanalytischen Settings (wie Liegen auf der Couch, mehrfach wöchentliche Therapiesitzungen) auf die Ergebnisse von psychoanalytischer Psychotherapie.
- 3) Der Einfluss der wöchentlichen Sitzungsfrequenz auf das Zustandekommen eines „analytischen Prozesses“ und ob bzw. inwieweit der Prozess in psychoanalytischer Psychotherapie als „fortgesetzte, zeitlich nicht befristete Fokalthherapie mit wechselnden Fokus konzeptualisiert werden kann“ (Thomä u. Kächele, zit. nach Hartkamp, 1997, S. 124).
- 4) Das Verhältnis von spezifischen, für psychoanalytische Psychotherapien geltenden Faktoren, wie Deutung und Interpretation in Relation zu unspezifischen emotional erfahrenen Interaktions- und Beziehungsfaktoren.

## Das Studiendesign zukünftiger Psychotherapieforschung

Zukünftige Forschung wird sich dem klinischen Alltag zuwenden, sowohl in privaten Praxen als auch in stationären Institutionen. Wie auch schon die Berliner Psychotherapiestudie (Rudolf, 1988), die in den frühen 80er Jahren konzipiert wurde, wird die zukünftige Forschung bemüht sein, psychoanalytische Therapie und Psychotherapie unter den realistischen Bedingungen der Alltagsarbeit zu evaluieren. Wünschenswert sind Aussagen aus der therapeutischen Praxis direkt für diese Praxis (Stuhr, 2001b).

Die Zeiten von „Wirksamkeitsstudien“, in denen 62 % aller untersuchten „Patienten“ freiwillige oder angeworbene Versuchspersonen, im Durchschnitt 23 Jahre alt waren, die Untersuchungen an highschools oder colleges durchgeführt wurden und die meisten „Therapeuten“ Doktoranden mit einer mittleren therapeutischen Erfahrung von drei Jahren (Parloff, 1982, zit. bei Bozock & Bühler, 1988), dürften der Vergangenheit angehören.

Allerdings erscheint auch der „Goldstandard“ klinischer Ergebnisstudien – prospektiver, doppelblinder, randomisierter, manualisierter Behandlungsplan (Fisher u. Greenberg, 1996) in der Ergebnisforschung bei Psychoanalysen kaum realisierbar und wegen seiner Praxisferne auch nicht erstrebenswert. Randomisierte Erfolgsstudien werden von Rüger u. Senf, 1994, sowie von Rudolf et al. 1994 für Kontrollgruppendesigns auch aus ethischen Gründen abgelehnt.

Es gibt sogar Autoren die zur Diskussion stellen, ob die Wirksamkeit von analytischer Psychotherapie überhaupt wissenschaftlich exakt messbar ist (Fisher u. Greenberg, 1996), und wer die Ergebnisse der Psychotherapie beurteilen soll. Seligman (1995) hält empirische Effizienzstudien sowieso nicht für das Maß aller Dinge, wenn es um die Frage geht, was in der Praxis wirkt.

### *Methodische Anforderungen*

1) Kontrollgruppen sind aus methodischen Gesichtspunkten dann sinnvoll, wenn

- die beiden verglichenen Gruppen hinsichtlich einer einzigen, sehr aussagekräftigen Variablen verglichen werden können,
- die Größe der verglichenen Gruppen so ist, dass sich Störvariablen aufheben und schließlich
- die Kontrollgruppe sich nicht in einem ungeprüften, aber wesentlichen Kriterium von der behandelten Gruppe unterscheidet.

Zu empfehlen sind „natürliche“ Kontrollgruppen, wo indizierte Patienten ihre Behandlung nicht antreten (z.B. durch Wegzug, Stuhr, 2001b). Wartekontrollgruppen sind den Kontrollgruppen ohne Behandlungsaussichten vorzuziehen.

Stuhr (2001b) hält Kontrollgruppenstudien nur zu Beginn der Einführung eines neuen Psychotherapieverfahrens für notwendig. Nach der Meinung von Rüger und Senf (1994) ist der klassische experimentelle Ansatz im Bereich der Ergebnisforschung kaum durchführbar, zumindest was mittlere und längere Behandlungszeiträume angeht. Die Patienten lassen sich in diesen Fällen nicht zufallsorientiert einer Behandlungsgruppe und einer Kontrollgruppe zuordnen.

Alternativ zum klassischen Kontrollgruppen-Design kann auch ein Vergleichsgruppen-Design gewählt werden, in dem unterschiedlich behandelte Patienten-Populationen bezüglich vorher festgelegter Ergeb-

nis-Kriterien geprüft werden. Wird eine Zufallsordnung vorgenommen, treten die Schwierigkeiten, wie beim klassischen Kontrollgruppen-Design hier in abgeschwächter Form auf. Wird keine Zufallszuordnung vorgenommen, besteht die Gefahr, dass sich die zu vergleichenden Gruppen in dem entscheidenden Ergebnis-Kriterium unterscheiden.

2) Multidimensionale Erfassung des Therapieergebnisses mit objektiven und subjektiven Daten (Rüger und Senf, 1994; so auch Langenmayr u. Kosfelder, 1998).

Der Behandlungserfolg lässt sich deshalb nur multidimensional erfassen, weil es so schwierig ist zu definieren, was eigentlich den Erfolg einer Behandlung ausmacht.

Die direkten Erfolgsbeurteilungen (durch die Patienten selbst) scheinen immer etwas höher zu liegen als die indirekten (mit vorher-/nachher Messungen), und Selbsteinschätzungen fallen häufig besser aus als Fremdbeurteilungen.

Für die quantitative Erfassung von Strukturveränderung durch psychoanalytische Psychotherapie sollen entsprechende Testverfahren entwickelt werden.

3) Intrapsychische und interpersonale Befunde sind auf unterschiedlichen Ebenen der Wahrnehmung zu dokumentieren. Es sind dies die folgenden Ebenen:

- a. Ebene des Expertenurteils (Therapeut sieht Patienten),
- b. Ebene der therapeutischen Beziehung (Therapeut sieht therapeutische Zusammenarbeit),
- c. Ebene der Gegenübertragung (Therapeut erlebt seine Reaktion auf den Patienten),
- d. Ebene der Selbsteinschätzung (Therapeut sieht sich selbst),
- e. Ebene der Selbsteinschätzung (Patient sieht sich selbst),
- f. Ebene der therapeutischen Beziehung (Patient sieht therapeutische Zusammenarbeit),
- g. Ebene der Übertragung (Patient sieht Therapeut) (s. Rudolf, 1988a),
- h. Intersubjektive Ebene (Teilnehmender Beobachter beobachtet Therapeut-Patient-Dyade).

4) Da es in allen ernsthaften Psychotherapien um das Verstehen des individuellen Leidens eines einzigartigen Menschen und seiner Geschichte geht, muss in einer dem Gegenstand angemessen empirisch-psychoanalytischen Forschung die „subjektive Einmaligkeit des Individuums“ Berücksichtigung finden (Leuzinger-Bohleber, 1997, so auch Stuhr, 2001a, S.141).

5) Qualitative methodische Ansätze sind neben die üblichen quantitativen Methoden zu stellen, bei denen in adäquater, differenzierter Weise sowohl der Einzelfall als auch der stattgefundenen psychotherapeutischen Prozess untersucht wird, was jedoch zeitaufwendig und anspruchsvoll ist.

#### *Anforderungen an die Katamnese:*

Leuzinger-Bohleber (1997) verweist auf die komplexen Forschungsprobleme der psychoanalytischen Katamneseforschung, die sich aus der spezifischen wissenschaftstheoretischen Position der Psychoanalyse erklärt, bei der es im Unterschied zur Verhaltenstherapie mit ihrem Ziel der Symptombefreiung dem

Analysanden überlassen ist, seine Ziele selbst bzw. überhaupt Ziele für sich zu bestimmen. Leuzinger-Bohleber resümiert, dass aus psychoanalytischer Sicht ein prinzipielles Dilemma unauflösbar ist, „einerseits die Idiosynkrasie des Einzelfalls in seiner Unvergleichbarkeit zu verstehen und andererseits diese Idiosynkrasie immer mit anderen zu vergleichen, d.h. sie in einen breiteren Zusammenhang zu stellen“ (Leuzinger-Bohleber, 1997, S.136).

Zur Problematik der Stabilität von Ergebnissen aus Katamneseuntersuchungen äußert sich Sandell (1999) und bezieht sich dabei auf die Ergebnisse von Smith et al. (1980), der zwischen der Effektstärke und dem Nachuntersuchungszeitraum bis zu zwei Jahren eine starke negative Korrelation fand, auf Kazdin (1994), der zeigte, dass sowohl die absoluten als auch die relativen Effekte der Behandlung im Verlauf der Zeit schwanken können und auf Shapiro et al. (1995), der bei der Nachuntersuchung nach drei Monaten bzw. nach einem Jahr zu ganz unterschiedlichen Langzeitergebniseinschätzungen kommt, je nachdem, in welcher Form und wie lange Therapien durchgeführt werden. Sandell (1999) betont, dass das Ergebnis selbst einen Prozess darstellt, das heißt, dass sich die Ergebnisbeurteilung nach Beendigung der Therapie mit der Zeit ändert, bei den Patienten mehr oder weniger autonome Prozesse ablaufen, die während der Therapie angestoßen und in Gang gesetzt worden sind. Heisterkamp (2001) führt diese Entwicklung auf die Selbstbehandlung als organisierendes Prinzip zurück, das Neues kreiert. Der Prozess nach Beendigung der Therapie kann genauso wechselhaft sein wie der Psychotherapie-Prozess selbst: es kann Hochs und Tiefs geben, Hoffnung und Verzweiflung usw.

Katamnestische Befragungen über mehrere Jahre sind notwendig, um die Dauerhaftigkeit von Behandlungsergebnissen zu untersuchen. Dabei sind Mehrpunktmessungen, d.h. Längsschnittverläufe statistischen Querschnittsproben vorzuziehen (Stuhr, 2001b).

Positive Nachwirkungen nach Therapieabschluss (psychoanalytische Langzeittherapien) berichten Rudolf et al. (1994) und von Rath et al. (1998). Neben den langfristigen Verbesserungen gibt es aber auch Verschlechterungen nach der Beendigung der (stationären) Behandlung, und das scheint nicht nur mit dem initialen Beschwerdedruck (niedriger initialer Beschwerdedruck = Verschlechterung u. u.) zu tun zu haben, sondern mit Umfeldbedingungen nach der Entlassung aus der Klinik, wie Partnerschaft oder Familie als Hintergrund, ambulante psychotherapeutische Weiterbehandlung, Arbeitsplatz mit stabilisierender Wirkung, die langfristige Verbesserungen begünstigen (Bassler, 1998). Deshalb sollten in der Katamneseerhebung auch Fragen nach dem beruflichen und sozialen Umfeld enthalten sein.

#### *Kriterien für die katamnestische Überprüfung von Behandlungsergebnissen:*

Wir beziehen uns hier im wesentlichen auf Rüger u. Senf (1994)

Vorab sind die Therapieziele festzulegen und eine Erfolgsdefinition unter Einschluss einer prognostischen Einschätzung.

Im einzelnen geht es dann um die folgenden Bereiche:

1) Beurteilung der Symptomatik: z.B. alte Symptomatik, andere psychische Symptome, sonstige Krankheiten, Arbeitsunfähigkeiten usw.).



- 2) Beurteilung der Änderung der Persönlichkeitsstruktur: z.B. Partnerbeziehung und Sexualität, berufliche Entwicklung, Kontaktverhalten.
- 3) Lebensentwicklung seit Abschluss der Behandlung (Bewährung und Bewältigung in Schwellen- und Krisensituationen).

Es sollten nicht nur objektive Daten (z.B. Krankenfehltag), standardisierte Daten (Tests, usw.), sondern auch subjektive Daten (z.B. Zufriedenheit) sowie interaktionelle Daten (Bewertung der therapeutischen Beziehung, sozialer Beziehungen, usw.) einbezogen werden. (Bedeutung multipler Perspektiven s. Kantrowitz, 1997, S.65). Dabei sind zusätzlich jeweils unterschiedliche Beobachterstandorte zu berücksichtigen, da Patienten, Therapeuten, unabhängige klinische Experten, Kostenträger usw. unterschiedliche Gewichtungen setzen. Durch ein solches Vorgehen können Gefahren vermieden werden wie beispielsweise

- durch eine eindimensionale Auswahl der Erfassungsinstrumente zu „richtigen“, aber wenig relevanten Ergebnissen zu kommen,
- wesentliche krankheitsrelevante Bereiche zu übersehen,
- subjektiv wichtig erlebte Veränderungen nicht ausreichend zu erfassen und zu werten und für den individuellen Verlauf bedeutsame Aspekte zu übersehen.

Ebenso sollen die wichtigsten Therapeutenvariablen (Geschlecht, Alter, Ausbildung, klinischer Erfahrungsstand etc.) erfasst werden.

Es sind mindestens Drei-Punkt-Untersuchungen (Behandlungsbeginn, Behandlungsabschluss, Katamnesezeitpunkt) notwendig mit einer ausreichend langen Katamnesezeitdauer.

Abschließend können wir feststellen, dass der Katamnesezeitpunkt in Zukunft ein bedeutsamer Stellenwert eingeräumt werden wird. Die Standards für ein methodisch einwandfreies, differenziertes Vorgehen dürften erheblich angehoben werden, um zu validen Ergebnissen zu kommen. Dies wird sowohl für die Grundlagenforschung, z.B. die differentielle Wirksamkeit von Therapiemethoden für Krankheitsgruppen als auch in der angewandten Psychotherapieforschung etwa zur Abklärung der Frage nach der Wirksamkeit unterschiedlicher Therapiedosen bei unterschiedlichen Patientenpopulationen von größter Bedeutung sein.

Für die psychoanalytische Wirkungsforschung sehen wir vier Themenkomplexe, die dringend weiterer Klärung bedürfen:

- 1) das wirksame Agens (der „Wirkstoff“) in der Beziehung zwischen Patient und Analytiker
- 2) die Indikation für Langzeitbehandlung
- 3) die Bedeutung der Frequenz für den Behandlungserfolg
- 4) Fragen des matchings von Analytiker und Patient

Neben möglichst repräsentativ angelegten praxisnahen Vergleichsstudien insbesondere für die Themen zwei bis vier ist für ein tiefenpsychologisches/psychodynamisches Verständnis der Zusammenhänge

mikroanalytische, in die Tiefe gehende qualitative Forschung und sicher dann auch Einzelfallforschung notwendig, um auf der Makroebene die Wirksamkeit bei bestimmten Subgruppen zu erhöhen.

In diesem Zusammenhang ist es notwendig, auch die Misserfolge (keine Veränderung oder Verschlechterung zum Behandlungsende und/oder im Katamnesezeitraum) genauestens zu analysieren. Was muss man als unabänderliche Tatsache hinnehmen, z.B. wenig motivierte Patienten haben geringere Erfolge, oder was kann der behandelnde Analytiker (an sich selbst?) tun, um einem schlecht motivierten Patienten zu mehr Motivation oder einem kontaktgehemmten zu mehr Offenheit zu verhelfen oder anders ausgedrückt, wie kann der Analytiker den „Wirkstoff“ verändern, damit seine Behandlungen in diesen (für ihn) schwierigen Fällen effektiver werden?

Sicherlich wäre es ein Trugschluss zu glauben, Psychotherapieforschung könnte jemals endgültige Antworten geben. Genau so wenig gibt es die „ideale Evaluationsstudie“ (Langenmayr u. Kosfelder, 1998, S. 107). Fragestellungen, Forschungsanliegen, Studiendesigns, Ergebnisinterpretationen, Schlussfolgerungen unterliegen wie auch die psychoanalytische Behandlungspraxis selbst dem ständig sich aufs Neue organisierenden und mehr oder weniger wechselvollen Geschehen zwischen allen an der Forschung Beteiligten.

Psychoanalytische Behandlung und Forschung ist und bleibt in Zukunft um die ständige Annäherung an die Wahrheit – „sub specie aeternitatis“ (A. Adler) – bemüht, wohlwissend, dass es viele, einem ständigen Wandlungsprozess unterliegende Wahrheiten gibt.

Was uns antreibt ist, Erkenntnis und Verstehen zu mehren, um unseren Patienten noch besser helfen zu können, ob als Online- oder Offline-Forscher (Moser). Wir hoffen, dass die vorliegende Arbeit einen, wenn auch nur bescheidenen Beitrag, dazu leisten konnte.

## 14. Literaturverzeichnis

Adorno, Th. W. (1972): Zur Logik der Sozialwissenschaften. In: Adorno, Th. W., Dahrendorf, R., Pilot, H., Albert, H., Habermas, J., Popper, K.R.: Der Positivismusstreit in der deutschen Soziologie. Neuwied u.Berlin (H. Luchterhand Verlag).

Bachrach, H. M., Weber, J. J., Solomon, M. (1985): Factors associated with the outcome of psychoanalysis (clinical and methodological considerations):Report of the Columbia Psychoanalytic Center Project(IV).International Review of Psychoanalysis 12, S. 79-88.

Bachrach, H. M., Galatzer-Levy, R., Skolnikoff, A., Waldron, J. S. (1991): On the efficacy of psychoanalysis. Journal of the American Psychoanalysis Association 39, S. 871-916.

Bassler, M., Potratz, B., Krauthauser, H. (1995a): Der „Helping Alliance Questionnaire“ (HAQ) von Luborsky. Psychotherapeut 40, S. 23-32.

Bassler, M., Krauthauser, H., Hoffmann, S. O. (1995b): Welche Faktoren beeinflussen die Dauer von stationärer Psychotherapie? Psychotherapie Psychosomatik medizinische Psychologie 45, S.167-175.

Bassler, M. (1998) (Hrsg.) : Wirkfaktoren von stationärer Psychotherapie. Mainzer Werkstatt über empirische Forschung von stationärer Psychotherapie. Gießen (Psychosozial-Verlag).

Benjamin, L. S. (1988): SASB short form user's Manual & Basic Principles of the Intrex.

Bergin, A. E. (1971):The evaluation of therapeutic outcome. In: Bergin, A. E., Garfield, S. L. (eds.): Handbook of psychotherapy and behavior change. New York (John Wiley & Sons), S. 217-270.

Bergin, A. E. u. Lambert, M. J. (1978):The evaluation of therapeutic outcomes. In: Garfield, S. L., Bergin, A. E. (eds.): Handbook of psychotherapy and behavior change, 2.Auflage. New York (John Wiley & Sons), S. 139-189.

Bergin, A. E. u. Garfield, S. L. (1994): Introduction and historical overview. In: Bergin, A. E., Garfield, S. L. (eds.): Handbook of psychotherapy and behavior change. New York (John Wiley & Sons), 4. Auflage, S. 3-18.

Böhme, H., Teusch, L. (1999): Persönlichkeitsfragebögen – die Auswirkung der Testabwehr in Persönlichkeitsfragebogen auf die Evaluation von Psychotherapieergebnissen. Psychotherapeut 44, S. 36-43.

Bozok, B. u. Bühler, K.-E. (1988): Wirkfaktoren der Psychotherapie – spezifische und unspezifische Einflüsse. Fortschr. Neurol. Psychiat. 56, S. 119-132.

Brähler, E., Scheer, J. W. (1995): Der Gießener Beschwerdebogen (GBB). Bern (H. Huber), 2. Auflage.

Brickenkamp, R. (Hrsg.) (1997): Handbuch psychologischer und pädagogischer Tests. Göttingen (Hogrefe), 2. Auflage.

Brockmann, J. (1995): Liefert die empirische Psychotherapieforschung relevante Ergebnisse für die Praxis des Psychoanalytikers? Forum der Psychoanalyse 11, S. 348-364.

Brockmann, J., Schlüter, T., Eckert, J. (2001): Die Frankfurt-Hamburg-Langzeit-Psychotherapiestudie – Ergebnisse der Untersuchung psychoanalytisch orientierter und verhaltenstherapeutischer Langzeit-Psychotherapien in der Praxis niedergelassener Psychotherapeuten. In: Stühr, U., Leuzinger-Bohleber, M., Beutel, M.: Langzeit-Psychotherapie. Stuttgart Berlin Köln (Kohlhammer), S. 271-276.

Broda, M., Lamberti, J., Dinger-Broda, A., Klinkenberg, N. (1999) : Entwicklung eines Fragebogens zur Standardanamnese für die stationäre psychosomatische Rehabilitation. Psychotherapie Psychosomatik medizinische Psychologie 49, S. 337-342.

Brunner, R. (1994): Überlegungen zur Entwicklung individualpsychologischer Psychotherapieforschung. Zeitschrift für Individualpsychologie 19, S. 168-190.

Burbiel, I., Apfelthaler, R., Fabian, E., Schanne, U., Seidler, G., Wolfrum, G. (1990): Die Effizienz stationärer Psychosentherapie in der Dynamisch-Psychiatrischen Klinik Mentschwaige München. Dynamische Psychiatrie 23, S. 247-301.

Burbiel, I., Apfelthaler, R., Fabian, E., Schanne, U., Wolfrum, G. (1992): Stationäre Psychotherapie der Psychosen. Eine testpsychologische katamnestiche Untersuchung. Dynamische Psychiatrie 25, S. 214-27.

Burbiel, I., Apfelthaler, R., Sandermann, G. (1994): Stationäre Psychotherapie von Borderline-Patienten. Eine Prä/Post- und katamnestische Studie. *Dynamische Psychiatrie* 27, S. 336-350.

Butler, S. F. u. Strupp, H. H. (1986): Specific and nonspecific factors in psychotherapy. A problematic paradigm for psychotherapy research. *Psychotherapy* 1, S. 30-40.

Clauß, G., Finse, F.-R., Partzsch, L. (1999): Statistik. Frankfurt a.M., 3. Auflage.

Dahrendorf, Ralf (1972): Anmerkungen zur Diskussion. In: Adorno, Th. W., Dahrendorf, R., Pilot, H., Albert, H., Habermas, J., Popper, K. R.: *Der Positivismusstreit in der deutschen Soziologie*. Neuwied u. Berlin (H. Luchterhand Verlag).

Davies-Osterkamp, S., Hartkamp, N., Junkert, B. (1993): Die Intrex-Kurzform. In: Tress, W. (Hrsg.): *Die Strukturelle Analyse Sozialen Verhaltens-SASB*. Heidelberg (R. Asanger Verlag), S. 156-219.

Derogatis, L. R., Lippman, R. S., Covi, L. (1973): SCL-90: An outpatient psychiatric rating scale - preliminary report. *Psychopharmacology Bulletin* 9, S. 13-28.

Dilling, H., Mombour, W., Schmidt, M. H. (Hrsg.) (1993): *Internationale Klassifikation psychischer Störungen ICD-10 Kapitel V(F)*. Bern, 2. Auflage.

Dührssen, A. (1962): Katamnestische Ergebnisse bei 1004 Patienten nach analytischer Psychotherapie. *Zeitschrift für Psychosomatische Medizin* 8, S. 94-113.

Dührssen, A., Jorswieck, E. (1965): Eine empirisch-statistische Untersuchung zur Leistungsfähigkeit psychoanalytischer Behandlung. *Nervenarzt* 36, S. 166-169.

Egle, U. T., Heucher, K., Hoffmann, S. O., Porsch, U. (1992): Psychoanalytisch orientierte Gruppentherapie mit psychogenen Schmerzpatienten: Ein Beitrag zur Behandlungsmethodik. *Psychotherapie Psychosomatik medizinische Psychologie* 42, S. 79-90.

Eysenck, H.-J. (1952): The Effects of Psychotherapy: An Evaluation. *Journal of Consulting Psychology* 16, S. 319-324.

Faller, H. (1994): Das Forschungsprogramm „Qualitative Psychotherapieforschung“. Versuch einer Standortbestimmung. In: Faller, H., Frommer, J. (Hrsg.): Qualitative Psychotherapieforschung. Heidelberg (R. Asanger), S. 15-37.

Faller, H., Frommer, J. (Hrsg.) (1994): Qualitative Psychotherapieforschung. Heidelberg (R. Asanger).

Fischer, G., Klein, B. (1997): Psychotherapieforschung. Forschungsepochen, Zukunftsperspektiven und Umriss eines dynamisch-behavioralen Verfahrens. In: Hildemann, K. D., Potthoff, P. (Hrsg.): Psychotherapie - Quo vadis? Göttingen (Hogrefe-Verlag), S. 17-35.

Fisher, S., Greenberg, R. P. (1996): Freud scientifically reappraised. New York (Wiley).

Fonagay, P., Kächele, H., Krause, R., Jones, E., Perron, R. (eds.) (1999): An Open Door Review of Outcome Studies in Psychoanalysis. London. International Psychoanalytic Association. <http://www.ipa.org.uk>

Franz, M., Schmitz, N., Lieberz, K., Schepank, H. (1998): Das multiple somatoforme Syndrom in der Allgemeinbevölkerung. In: Rudolf, G. u. Henningsen, P. (Hrsg.): Somatoforme Störungen. Stuttgart (Schattauer), S. 41-51.

Franz, M., Häfner, S., Lieberz, K., Reister, G., Tress, W. (2000): Spontanverlauf psychogener Beeinträchtigung in einer Bevölkerungsstichprobe über elf Jahre. Psychotherapeut 45, S. 100-107.

Gehring, A., Blaser, A. (1983): MMPI, Deutsche Kurzform für Handauswertung, Handbuch. Bern (H. Huber).

Goudsmit, A. L. (ed.) (1989): Self Organization in Psychotherapy. New York (Springer).

Grande, T., Porsch, U., Rudolf, G. (1988): Muster therapeutischer Zusammenarbeit und ihre Beziehung zum Therapieergebnis. Zeitschrift für Psychosomatische Medizin 34, S. 76-100.

Grande, T., Rudolf, G., Oberbracht, C. (1997): Die Praxisstudie Analytische Langzeittherapie. Ein Projekt zur prospektiven Untersuchung. In: Leuzinger-Bohleber, M., Stühr, U.: Psychoanalysen im Rückblick, Gießen (Psychosozial-Verlag), S. 415-431.

Grawe, K. (1988): Zurück zur psychotherapeutischen Einzelfallforschung. Zeitschrift für Klinische Psychologie 17, S. 1-7.

Grawe, K. (1992): Psychotherapie zu Beginn der neunziger Jahre. Psychologische Rundschau 43, S. 132-162.

Grawe, K., Bernauer, F. u. Donati, R. (1990): Psychotherapien im Vergleich. Haben wirklich alle einen Preis verdient? Psychotherapie Psychosomatik medizinische Psychologie 40, S. 102-114.

Grawe, K., Donati, R. u. Bernauer, F. (1994): Psychotherapie im Wandel. Von der Konfession zur Profession. Göttingen (Hogrefe).

Greif-Higer, G., Egle, U. T. (1993): Chronischer Schmerz im Thoraxbereich In: Egle, U. T., Hoffmann, S. O.: Der Schmerzkranken. Stuttgart, New York (Schattauer), S. 511-529.

Groß, A., Riedel, W. P. (1995): Therapieergebnis und Komplementarität in der Therapeut-Patient-Beziehung. Regensburg (S. Roderer Verlag).

Habermas, J. (1972): Analytische Wissenschaftstheorie und Dialektik. In: Adorno, Th. W., Dahrendorf, R., Pilot, H., Albert, H., Habermas, J., Popper, K. R.: Der Positivismusstreit in der deutschen Soziologie. Neuwied u. Berlin (H. Luchterhand Verlag).

Habermas, J. (1973): Erkenntnis und Interesse. Frankfurt a.M..

Habermas, J. (1982): Zur Logik der Sozialwissenschaften. Frankfurt a.M..

Hampe, M. (2001): Theorie, Erfahrung, Therapie. Anmerkungen zur philosophischen Beurteilung psychoanalytischer Prozesse. Psyche 55, S. 328-336.

Hartkamp, N. (1997): Psychoanalytische Therapie: Ergebnisse und Prozesse – Was wissen wir und wonach müssen wir fragen? In: Tschuschke, V., Heckrath, C., Tress, W. (Hg.) (1997): Zwischen Konfusion und Makulatur. Zum Wert der Berner Metaanalyse von Grawe et al. Göttingen (Vandenhoeck & Ruprecht), S. 106-124.

Hathaway, S. R. u. McKinley, J. C. (1991): MMPI Saarbrücken. Bern (H. Huber), 3. Nachdruck.

Heckrath, C., Dohmen, P.(1994): Wege und Umwege der Psychotherapieforschung. Zeitschrift für Individualpsychologie 19, S. 156-167

Heinzel, R., Breyer ,F., Klein ,T. (1996): Ambulante Psychoanalyse in Deutschland. Eine katanestische Evaluationsstudie. Fakultät für Wirtschaftswissenschaften und Statistik, Universität Konstanz.

Heisterkamp, G. (1999a): Zur Freude in der analytischen Psychotherapie. Psyche 53, S. 1247-1265.

Heisterkamp, G. (1999b): Indizes fehlender Freude in der Tiefenpsychologie. Zeitschrift für Individualpsychologie 24, S. 381-402.

Heisterkamp, G. (2001): Mittelbares und unmittelbares Verstehen im psychotherapeutischen Handlungsdialog. In: Milch, W. E., Wirth, H.-J. (Hg.): Psychosomatik und Kleinkindforschung. Gießen (Psychosozial-Verlag).

Henry, W. P., Strupp, H. H., Schacht, T. E., Gaston, L. (1994): Psychodynamic Approaches. In: Bergin, A. E., Garfield, S. L. (eds.). Handbook of psychotherapy and behavior change. New York (John Wiley & Sons), 4. Auflage. S. 467-508.

Heuft, G., Seibüchler-Engel, H., Taschke, M., Senf, W. (1996): Langzeitoutcome ambulanter psychoanalytischer Psychotherapien und Psychoanalysen. Forum der Psychoanalyse 12, S. 342-355.

Huber, D., Kächele, H. (2000): Die „Internationale“ der psychodynamischen Psychotherapieforscher. Psychotherapie Psychosomatik medizinische Psychologie 50, S.309

Hutterer, A. (1996): Die Consumer Reports Studie: Längere Psychotherapien sind effektiver! Psychotherapie Forum 4, Nr.1, Supplement, S. 2-6.

Kächele, H.(1990): Wie lange dauert Psychotherapie? Psychotherapie Psychosomatik medizinische Psychologie 40, S.148-151.

Kächele, H. (1992): Psychoanalytische Therapieforschung 1930-1990. In: Psychoanalyse-Klinik und Kulturkritik (Psyche) 46, S. 259-279.



Kächele, H. (1995): Klaus Grawes Konfession und die psychoanalytische Profession. *Psyche* 49, S. 481-492.

Kaiser, E. (1993): Quantitative Psychotherapieforschung – modernes Paradigma oder Potemkinsches Dorf. *Forum der Psychoanalyse* 9, S. 348-366.

Kaiser, E. (1995): Der psychotherapeutische Weltgeist zu Bern. *Psyche* 49, S. 493-513.

Kammerer, H. (1998): Stationäre Psychotherapie. Eine empirische Untersuchung mit der Strukturalen Analyse Sozialen Verhaltens (SASB). Dissertation zur Erlangung des Grades eines Doktors der Medizin/Zahnmedizin der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf

Kantrowitz, J. L. (1997): Ergebnisforschung in der Psychoanalyse: Rückblick und Neubewertung. In: Leuzinger-Bohleber, M. u. Stuhr, U. (Hg.): *Psychoanalysen im Rückblick*. Gießen (Psychosozial-Verlag). S.61-72.

Kazdin, A. E. (1994): Methodology, design and evaluation in psychotherapy research. In: Bergin, A. E. u. Garfield, S. L. (eds.): *Handbook of psychotherapy and behavior change*. New York (John Wiley & Sons), 4. Auflage, S. 19-71.

Keller, G., Westhoff, G., Dilg, R., Rohner, R., Studt, H. H. (1997): Wirksamkeit und Kosten-Nutzen-Aspekte ambulanter jungianischer Psychoanalysen und Psychotherapien - eine Katanese-Studie. Berlin.

Kern, C., Küchenhoff, J. (2000): Zur Evaluation der Psychotherapie im teilstationären Setting. In: Bassler, M. (Hrsg.): *Wirkfaktoren von stationärer Psychotherapie/Mainzer Werkstatt über empirische Forschung*. Gießen (Psychosozial-Verlag), S. 24-55.

Kernberg, O. F. (1997): Widerstände gegen Forschung in der Psychoanalyse. In: Leuzinger-Bohleber, M. u. Stuhr, U. (Hg.): *Psychoanalysen im Rückblick*. Gießen (Psychosozial-Verlag), S. 39-45.

Klein, D. F., Godwin Rabkin, J. (1993): Specify and strategy in psychotherapy research and practice. In :Miller, N.E., Luborsky, L., Barber, J.P., Docherty, J. P. (eds): *Psychodynamic treatment research – a handbook for clinical practice*. New York (Basic Books).

Kordy, H., von Rad, M., Senf, W. (1988): Die Zeit und ihre Bedeutung für eine erfolgreiche Psychotherapie. *Psychotherapie Psychosomatik medizinische Psychologie* 39, S. 212-222.

Kordy, H., von Rad, M., Senf, W. (1989): Empirische Hypothesen über die psychotherapeutische Behandlung psychosomatischer Patienten in Kurz- und Langzeittherapie. *Praxis Psychotherapie und Psychosomatik* 52, S. 155-163.

Kordy, H., von Rad, M., Senf, W. (1990): Therapeutische Faktoren bei stationärer Psychotherapie: Die Sicht der Patienten. *Psychotherapie Psychosomatik medizinische Psychologie* 40, S. 380-387.

Kordy, H., Kächele, H. (1996): Ergebnisforschung in Psychotherapie und Psychosomatik. In: Adler, R. H., Hermann, J. M., Köhler, K., Schonecke, O. W., von Uexküll, T., Wesiak, W. (Hg.): *Psychosomatische Medizin*. München (Urban und Schwarzenberg), S. 490-501.

Kubinger, K. (1999): Forschung in der psychologischen Diagnostik. *Psychologische Rundschau* 50, S. 131-139.

Lambert, M. J. (1992): Psychotherapy outcome research: Implications for integrative and eclectic therapists. In: Norcross, G. C. u. Goldfried, M. R. (eds.): *Handbook of psychotherapy integration*. New York (Basic Books).

Lambert, M. J., Bergin, A. E. (1994): The effectiveness of psychotherapy. In: Bergin, A. E. u. Garfield, S. L.: *Handbook of psychotherapy and behavior change*. New York (J. Wiley & Sons), 4. Auflage, S. 143-189.

Lambert, M. J. u. Hill, C. E. (1994): Assessing psychotherapy outcomes and processes. In: Bergin, A. E. u. Garfield, S. L.: *Handbook of psychotherapy and behavior change*. New York (J. Wiley & Sons), 4. Auflage, S. 72-113.

Langenmayr, A., Kosfelder, J. (1998): Evaluation in der Psychotherapie. In: Menne, K. (Hg.): *Qualität in Beratung und Therapie*. Bundeskonferenz für Erziehungsberatung. Weinheim und München (Juventa Verlag), S. 95-114.

Leikert, S., Ruff, W. (1994): Methodische Prinzipien psychoanalytischer Therapieforschung. *Forum der Psychoanalyse* 10, S. 77-86.

Leimkühler, A. M., Müller, U. (1996): Patientenzufriedenheit-Artefakt oder soziale Tatsache? *Nervenarzt* 67, S. 765-773.

Leuzinger-Bohleber, M. (1994): Die Einzelfallstudie als psychoanalytisches Forschungsinstrument. *Psyche* 49, S. 434-480.

Leuzinger-Bohleber, M. (1997): Psychoanalytische Katamneseforschung. In: Leuzinger-Bohleber, M. u. Stuhr, U. (Hg.): Psychoanalysen im Rückblick. Gießen (Psychosozial-Verlag), S. 125-163.

Leuzinger-Bohleber, M., Kächele, H. (1990): Von Calvin zu Freud: Fünf aggregierte Einzelfallstudien zur Veränderung kognitiver Prozesse in Psychoanalysen. Zeitschrift für Klinische Psychologie 19, S. 111-122.

Leuzinger-Bohleber, M., Stuhr, U., Beutel, R. (2001): Langzeitwirkungen von Psychoanalysen und Psychotherapien: Eine multiperspektivische, repräsentative Katamnese studie. Psyche 55, S. 193-276.

Luborsky, L., Singer, B. u. Luborsky, L. (1975): Comparative studies of psychotherapies. Archives of General Psychiatry 32, S. 995-1008.

Marziali, E., Marmar, C. u. Krupnick, J. (1981): Therapeutic alliance scales: Development and relationship to psychotherapy outcome. American Journal of Psychiatry 138, S. 361-364.

Mayntz, R., Holm, K., Hübner, P. (1972): Einführung in die Methoden der empirischen Soziologie. Opladen, 3. Auflage.

McNeilly, C. L. u. Howard, K. I. (1991): The effects of psychotherapy: A reevaluation based on dosage. Psychotherapy Research 1, S. 74-78.

Meltzoff, J., Kornreich, M. (1970): Research in psychotherapy. New York (Atherton Press).

Mertens, W. (1994): Psychoanalyse auf dem Prüfstand? Eine Erwiderung auf die Meta-Analyse von Klaus Grawe. Berlin-München (Quintessenz).

Mertens, W. (1995): Warum Psychoanalysen lange dauern. Psyche 49, S. 405-433.

Meyer, A.-E. (1990): Eine Taxonomie der bisherigen Psychotherapieforschung. Editorial. Zeitschrift für klinische Psychologie 19, S. 287-291.

Meyer, A.-E., Richter, R., Grawe, K., Graf v.d.Schulenberg, J.-M., Schulte, B. (1991): Forschungsgutachten zu Fragen eines Psychotherapeutengesetzes. Universitätskrankenhaus Hamburg-Eppendorf.

Orlinsky, D. E. u. Howard, K. J. (1986): Process and outcome in psychotherapy. In: Garfield, S. L. u. Bergin, A. E. (eds.): Handbook of psychotherapy and behavior change. New York (J.Wiley), 3. Auflage, S. 311-384.

Paar, G., H., Kriebel, R., Noack, R. (1994): Eine empirische Analyse zum Therapieabbruch bei stationärer Psychotherapie in einer Psychosomatischen Fachklinik. In: Hahn, P., Werner, A., Bergmann, G., Drinkmann, A., Eich, W., Hayden, M., Herzog, W.: Modell und Methode in der Psychosomatik. Eine Bestandsaufnahme gegenwärtiger psychosomatischer Forschungs- und Arbeitsansätze. Weinheim (Deutscher Studien Verlag), S. 292-296.

Parloff, M. B. (1982): Psychotherapy research evidence and reimbursement decisions: Bambi meets Godzilla. American Journal of Psychiatry 139, S. 718-727.

Popper, K., R. (1972): Die Logik der Sozialwissenschaften. In: Adorno, Th. W., Dahrendorf, R., Pilot, H., Albert, H., Habermas, J., Popper, K. R.: Der Positivismusstreit in der deutschen Soziologie. Neuwied.

Presslich-Titscher, E., Datler, W. (1994): Wirksamkeitsforschung in der Psychotherapie und öffentlicher Legitimationsdruck. Zeitschrift für Individualpsychologie 19, S. 203-213.

Remschmidt, H., Wienand, F., Wewetzer, C. (1990): Der Langzeitverlauf der Anorexia nervosa. In: Remschmidt, H., Schmidt, M., H.; Anorexia nervosa. Toronto (Hogrefe & Huber). Child and youth psychiatry: European perspectives 25, S. 127-136

Riedel, H., Schneider-Düker, M. (1991): Kontextbedingungen „kontrollierter“ und „unkontrollierter“ Psychotherapieforschung. Psychologische Rundschau 42, S. 19-28.

Rogner, J. (1994a): Ein Jahr nach Abschluss einer analytischen Adlerianischen Psychotherapie I: Vergleich mit Personen, die eine Psychotherapie beginnen. Zeitschrift für Individualpsychologie 19, S. 191-202.

Rogner, J. (1994b): Ein Jahr nach Abschluss einer analytischen Adlerianischen Psychotherapie II: Veränderungen der Symptomatik. Zeitschrift für Individualpsychologie 19, S. 318-326.

Rohrmeier, F. (1982): Langzeiterfolge psychosomatischer Therapien. Berlin.

Rudolf, G. (1991): Die therapeutische Arbeitsbeziehung. Berlin-Heidelberg (Springer-Verlag).

Rudolf, G., Grande, T., Porsch, U. (1988a): Die initiale Patient-Therapeut-Beziehung als Prädiktor des Behandlungsverlaufs. Zeitschrift für Psychosomatische Medizin 34, S. 2-18.

Rudolf, G., Grande, T., Porsch, U. (1988b): Die Berliner Psychotherapie-Studie. Zeitschrift für Psychosomatische Medizin 34, S. 32-49.

Rudolf, G., Manz, R. (1993): Zur prognostischen Bedeutung der therapeutischen Arbeitsbeziehung aus der Perspektive von Patienten und Therapeuten. Psychotherapie Psychosomatik medizinische Psychologie 43, S. 193-199.

Rudolf, G., Manz, R., Öri, Ch. (1994): Ergebnisse psychoanalytischer Therapien. Zeitschrift für Psychosomatische Medizin 40, S. 25–40.

Rüger, U., Senf, W. (1994): Evaluative Psychotherapieforschung: Klinische Bedeutung von Psychotherapie-Katamnesen. Zeitschrift für Psychosomatische Medizin 40, S. 103-116.

Sandell, R., Blomberg, J., Lazar, A., Schubert, J., Carlsson, J., Broberg, J. (1999): Wie die Zeit vergeht - Langzeitergebnisse von Psychoanalysen und analytischer Psychotherapie. Forum der Psychoanalyse 15, S. 327-347.

Sandell, R., Blomberg, J., Lazar, A., Carlsson, J., Broberg, J., Schubert, J. (2001): Unterschiedliche Langzeitergebnisse von Psychoanalysen und Langzeitpsychotherapien. Aus der Forschung des Stockholmer Psychoanalyse- und Psychotherapieprojekts. Psyche 55, S. 277-310.

Schauenburg, H., Strack, M. (1998): Die Symptom Checklist-90-R zur Darstellung von statistisch und klinisch signifikanten Psychotherapieergebnissen. Psychotherapie Psychosomatik medizinische Psychologie 48, S. 257-264.

Schulz, H., Lang, K., Lotz-Rambaldi, W., Bürger, W., Koch, U. (1999): Analyse von Behandlungsabbrüchen in der stationären psychosomatischen Rehabilitation anhand von Basisdokumentationen zweier Klinikträger. Psychotherapie Psychosomatik medizinische Psychologie 49, S. 326-336.

Seligmann, E. P. (1996): Die Effektivität von Psychotherapie – Die Consumer Reports-Studie. Integrative Therapie 2-3, S. 264-287.

Shapiro, D. A., Rees, A., Barkham, M., Hardy, G. (1995): Effects of treatment duration and severity of depression on the maintenance of gains after cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology* 63, S. 378-387.

Sloane, R. B., Stables, F. R., Cristol, A. H., Yorkstone, N. J., Whipple, K. (1975): *Psychotherapy versus behavior therapy*. Cambridge(Harvard Univ. Press).

Smith, M. L., Glass, G. V., Miller, T. I. (1980): *The benefits of psychotherapy*. Baltimore (J.Hopkins Univ. Press).

Strauss, B., Hess, H. (1993): Interpersonale Probleme, interpersonale Orientierung und Behandlungserfolg nach stationärer Gruppenpsychotherapie. *Psychotherapie Psychosomatik medizinische Psychologie* 43, S. 82-92.

Strupp, H. H., Hadley, S. W. u. Gomes-Schwartz, B. (1977): *Psychotherapy for better or worse: The problem of negative effects*. New York (J. Aronson).

Strupp, H. H. (1978): *Psychotherapy research and practice:an overview*. In: Garfield, S. L., Bergin, A. E.(eds.): *Handbook of psychotherapy and behavior change*. New York (J. Wiley and Sons), S. 3-22.

Strupp, H. H. u. Hadley, S. W. (1979): Specific versus nonspecific factors in psychotherapy. *Archives of General Psychiatry* 36, S. 1125-1136.

Stuhr, U. (2001a): Methodische Überlegungen zur Kombination qualitativer und quantitativer Verfahren in der psychoanalytischen Katamneseforschung und Hinweise zu ihrer Integration. In: Stuhr, M., Leuzinger-Bohleber, M., Beutel, M. E. (Hg.): *Langzeitpsychotherapie*. Stuttgart (Kohlhammer).

Tschuschke, V., Kächele, H., Hölzer, M. (1994): Gibt es unterschiedlich effektive Formen von Psychotherapie? *Psychotherapeut* 39, S.281-297.

Tschuschke, V., Heckrath, C., Tress, W. (1997): *Zwischen Konfusion und Makulatur. Zum Wert der Berner Metaanalyse von Grawe et al..* Göttingen (Vandenhoeck & Ruprecht).

Tschuschke, V., Bänninger-Huber, E., Faller, H., Fikentscher, E., Fischer, G., Frohburg, I., Hager, W., Schiffler, A., Lamprecht, F., Leichsenring, F., Leuzinger-Bohleber, M., Rudolph, G.,

Kächele, H. (1998): Psychotherapieforschung - wie man es (nicht) machen sollte. *Psychotherapie Psychosomatik medizinische Psychologie* 48, S. 430-444 .

Tschuschke, V. u. Anbeh, T. (2000): Wirkungen ambulanter analytischer Gruppenpsychotherapie in Deutschland. *Psychotherapeuten Forum* 7, S. 5–12.

v.Rad, M., Senf, W., Bräutigam, W.(1998): Psychotherapie und Psychoanalyse in der Krankenversorgung: Ergebnisse des Heidelberger Katamnese-Projektes. Psychotherapie Psychosomatik medizinische Psychologie 48, S. 88-100.

Vorberg, D., Blankenberger,S. (1999): Die Auswahl statistischer Tests und Maße. Psychologische Rundschau 50, S.157-164.

Wallerstein, R. S. (1990): Zum Verhältnis von Psychoanalyse und Psychotherapie. Wiederaufnahme einer Diskussion. Psyche 44, S. 967-994.

Wallerstein, R. S. (1997): Katamnesen in der Psychoanalyse: Zu ihrem klinischen und empirischen Wert. In: Leuzinger-Bohleber, M. u. Stuhr, U.: Psychoanalysen im Rückblick. Gießen (Psychosozial-Verlag), S. 46-60.

Weber, B. (1994): Über die Wirksamkeit der daseinsanalytischen Therapie bei psychosomatischen Störungen. Daseinsanalyse 11, S. 1-98.

Wilke, S. (1994): Einige Überlegungen zur Angemessenheit qualitativer Methoden für die Untersuchung psychoanalytischer Dialoge. In: Faller, H. u. Frommer, J. (Hg.): Qualitative Psychotherapieforschung. Heidelberg (R. Asanger), S. 73-93.

Wittmann, W. W., Matt, G. E. (1986): Meta-Analyse als Integration von Forschungsergebnissen am Beispiel deutschsprachiger Arbeiten zur Effektivität von Psychotherapie. Psychologische Rundschau 37, S. 20-40.

Wunderlich, G. (1993):Verringerung von Körperbeschwerden durch Psychotherapie. Praxis der Psychotherapie und Psychosomatik 38, S. 361-367.

Zielke, M. u. Kopf-Mehnert, C. (1973): Veränderungsfragebogen des Erlebens und Verhaltens, VEV. Weinheim (Beltz).

Zielke, M. (1993): Wirksamkeit stationärer Verhaltenstherapie.  
Psychosomatische Fachklinik, Bad Dürkheim.  
Weinheim (Psychologie Verlags Union).

# **ANHANG**

## **zur Studie**

**Die Wirksamkeit Individualpsychologisch-  
Psychoanalytischer Psychotherapie  
- eine Langzeitstudie zur Psychotherapieforschung**



## **Inhalt**

### **1. Tabellen**

- a. **Tabelle 23: MMPI(K)-T-Werte**
- b. **Tabellen 44 bis 79 b (Varianzanalysen)**
- c. **Tabelle 83 a – f (Soziodemografische Merkmale der Katamneseteilnehmer)**

- 2. Auswertung der Katamnesefrage: „Schildern Sie bitte möglichst konkret eine freudige Situation aus der Therapie.“**
- 3. Erstes Anschreiben an die 230 Psychoanalytiker/innen vom 15.03.1993 und Schreiben des Vorsitzenden der DGPT, Prof. Dr. Streeck, vom 18.09.1992**
- 4. Schreiben vom 30.06.1993 an alle mitwirkenden Psychoanalytiker/innen, incl. Untersuchungsdesign-Schema, Ablaufplan, Diagnoseschlüssel und Stellungnahme zu Bedenken**
- 5. Schreiben vom 30.06.1993 an alle nicht antwortenden Psychoanalytiker/innen**
- 6. Schreiben vom 21.04.1994 an alle mitwirkenden Psychoanalytiker/innen**
- 7. Schreiben vom September 1994 an alle mitwirkenden Psychoanalytiker/innen**
- 8. Schreiben vom Oktober 1994 an alle mitwirkenden Psychoanalytiker/innen**
- 9. Anschreiben an alle Patienten der Untersuchungsgruppe und der Kontrollgruppe zum Messzeitpunkt t1**
- 10. Anschreiben an alle Patienten der Untersuchungsgruppe und der Kontrollgruppe zum Messzeitpunkt t2, einem halben Jahr nach t1**
- 11. Anschreiben an alle Patienten der Untersuchungsgruppe und der Kontrollgruppe zum Messzeitpunkt t3, einem halben Jahr nach t2 und einem Jahr nach t1**
- 12. Anschreiben an alle Patienten der Untersuchungsgruppe zum Messzeitpunkt t4 und t5**
- 13. Schreiben vom 24.08.1999 an alle mitwirkenden Psychoanalytiker/innen zur Vorbereitung der Katamnese**

**14. Katamneseanschreiben an alle Psychoanalytiker/innen mit Anschreiben an die Patienten und mit Fragebögen zur Katamnese**

**15. Anschreiben zur Nachfassaktion im Sommer 2000**

## **1. Tabellen**

- **Tabelle 23: MMPI(K)-T-Werte**
- **Tabellen 44 bis 79 b (Varianzanalysen)**
- **Tabelle 83 a – f (Soziodemografische Merkmale der Katamneseteilnehmer)**

**Tabelle 23: MMPI(K): T-Werte für die 13 Skalen,  
Vergleich Untersuchungsgruppe zu t1 mit Kontrollgruppe zu t1**

	UG						KG					
	N	M Gesamt	N	M Männer	N	M Frauen	N	M Gesamt	N	M Männer	N	M Frauen
MMPI-Lie	142	45,30	42	44,67	100	45,57	50	47,02	14	49,00	36	46,25
MMPI-Validity	139	58,77	40	56,20	99	59,80	45	56,62	13	59,38	32	55,50
MMPI-Correction	140	47,21	42	46,95	98	47,32	45	46,80	14	46,28	31	47,03
MMPI-Hypochondriasis	141	60,64	39	60,77	102	60,59	50	60,36	14	67,93	36	57,42
MMPI-Depression	144	68,12	41	66,78	103	68,66	49	65,63	13	72,38	36	63,19
MMPI-Hysteria	137	66,94	39	66,36	98	67,17	48	67,02	13	73,69	35	64,54
MMPI-Psychopathie	140	63,33	40	69,30	100	60,95	49	58,14	14	69,14	35	53,74
MMPI-maskulin-feminin	136	65,13	38	66,68	98	64,54	42	60,69	12	64,75	30	59,06
MMPI-Paranoia	133	49,69	39	62,00	95	44,77	49	51,88	13	62,77	36	47,94
MMPI-Psychasthenia	134	60,87	40	58,15	95	61,09	48	58,42	14	60,71	34	57,47
MMPI-Schizoidie	133	60,57	40	60,72	93	60,50	47	57,81	13	67,54	34	54,09
MMPI-Hypomanie	134	56,54	36	55,03	98	56,50	46	56,67	13	59,07	33	55,27
MMPI-Soziale Introversion	141	60,93	41	63,78	100	59,76	49	59,08	14	66,36	35	56,17

### **Anmerkung zu den Tabellen der Varianzanalysen:**

Sie enthalten neben N (Anzahl der Patienten mit auswertbaren Antworten), M (Mittelwert) und Standardabweichung die Signifikanzen für den Cochrans-Test (Coch.) zur Überprüfung der Homogenität bzw. Heterogenität der Subgruppen (Voraussetzung für die Varianzanalyse) und die Signifikanz von F, des Unterschiedes zwischen den Subgruppen zum jeweils ersten Vergleichszeitpunkt, des Unterschiedes zwischen den jeweils zwei Zeitpunkten für die Gesamtgruppe (Signifikanz des Zeiteffektes) und der Signifikanz des Unterschiedes im Zeitverlauf vom ersten bis zum zweiten Zeitpunkt für die jeweiligen Subgruppen.

Da sich die Signifikanzen für die Gesamtgruppe in jeder Tabelle wiederholen, haben wir sie bei der Besprechung der Varianzanalysen nur vereinzelt erwähnt.

**Tabelle 44: Varianzanalyse (Manova): Diagnosegruppen "Neurotische Störung", "Affektive Störung", "Spezifische Persönlichkeitsstörung", ihre Entwicklung von t1 bis t3 in der Untersuchungsgruppe, gemessen mit den GBB-Skalen 1 bis 5 und den MMPI(K)-Skalen HS, D, HY und PT**

	Diagnose- gruppen		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	Coch. P	Coch. P	F	F	F
<b>GBB1-Erschöpfungsneigung</b>	Neurot. Störung	30	10,867	6,061	7,233	5,600	0,754	1,000	0,963	0,000	0,423
	Affektive Störung	50	10,960	5,678	6,920	5,074					
	Persönlichkeitsst.	28	10,429	5,364	8,071	5,484					
	Gesamt	108	10,796	5,660	7,306	5,301					
<b>GBB2-Magenbeschwerden</b>	Neurot. Störung	31	5,323	3,893	3,581	3,462	0,148	0,613	0,998	0,000	0,752
	Affektive Störung.	51	5,235	4,325	3,588	3,590					
	Persönlichkeitsst.	26	5,577	5,194	3,346	2,911					
	Gesamt	108	5,343	4,394	3,528	3,372					
<b>GBB3-Gliederschmerzen</b>	Neurot. Störung	28	8,143	5,205	5,750	4,592	0,284	0,294	0,836	0,000	0,941
	Affektive Störung	49	8,653	5,936	5,918	4,353					
	Persönlichkeitsst.	28	7,929	4,618	5,464	3,191					
	Gesamt	105	8,324	5,379	5,752	4,111					
<b>GBB4-Herzbeschwerden</b>	Neurot. Störung	29	6,655	4,353	4,931	4,780	0,053	0,003	0,094	0,000	0,883
	Affektive Störung	50	5,320	3,107	3,180	3,330					
	Persönlichkeitsst.	26	5,346	3,334	3,385	2,830					
	Gesamt	105	5,695	3,560	3,714	3,725					
<b>GBB5-Beschwerdedruck</b>	Neurot. Störung	26	29,462	13,932	20,346	13,667	1,000	0,487	0,982	0,000	0,744
	Affektive Störung	47	30,128	14,479	19,404	12,370					
	Persönlichkeitsst.	25	28,520	13,488	20,120	11,267					
	Gesamt	98	29,541	13,960	19,837	12,341					
<b>MMPI-Hypochondriasis</b>	Neurot. Störung	28	9,750	3,950	6,929	4,472	0,795	1,000	0,972	0,000	0,559
	Affektive Störung	50	9,360	4,707	7,320	4,705					
	Persönlichkeitsst.	27	9,407	4,643	7,704	4,496					
	Gesamt	109	9,476	4,462	7,314	4,556					
<b>MMPI-Depression</b>	Neurot. Störung	29	14,793	4,346	9,897	5,361	0,600	0,595	0,416	0,000	0,019
	Affektive Störung	50	13,680	4,547	10,620	5,379					
	Persönlichkeitsst.	28	14,357	3,714	12,536	4,114					
	Gesamt	107	14,159	4,276	10,925	5,128					
<b>MMPI-Hysteria</b>	Neurot. Störung	29	15,310	4,269	11,138	5,449	0,321	0,293	0,272	0,000	0,152
	Affektive Störung	43	15,209	5,608	11,837	6,451					
	Persönlichkeitsst.	27	15,963	5,042	14,370	5,212					
	Gesamt	99	15,444	5,053	12,323	5,933					
<b>MMPI-Psychasthenia</b>	Neurot. Störung	31	15,065	5,079	11,258	5,791	0,468	1,000	0,678	0,000	0,145
	Affektive Störung	47	14,681	5,402	11,660	5,738					
	Persönlichkeitsst.	26	14,846	4,240	13,423	5,308					
	Gesamt	104	14,837	4,993	11,981	5,660					

**Tabelle 45: Varianzanalyse (Manova): Diagnosegruppen**  
**"Neurotische Störung", "Affektive Störung", "Spezifische Persönlichkeitsstörung",**  
**ihre Entwicklung von t3 bis t5 in der Untersuchungsgruppe, gemessen mit den GBB-Skalen 1 bis 5,**  
**den MMPI(K)-Skalen HS, D, HY, PT und dem VEV**

	Diagnose- gruppen		t3		t5				Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes in Zeitverlauf
		N	M	s	M	s	Coch. P	Coch. P	F	F	F
GBB1-Erschöpfungsneigung	Neurot. Störung	15	8,400	6,029	9,400	4,437	0,601	0,876	0,276	0,441	0,203
	Affektive Störung	32	6,937	5,168	6,031	5,258					
	Persönlichkeitsst.	18	8,278	5,200	7,000	5,099					
	Gesamt	65	7,646	5,343	7,077	5,140					
GBB2-Magenbeschwerden	Neurot. Störung	15	4,133	3,482	3,533	2,475	0,753	0,070	0,497	0,152	0,946
	Affektive Störung	31	3,677	3,260	3,065	3,633					
	Persönlichkeitsst.	17	2,824	2,899	2,471	2,552					
	Gesamt	63	3,556	3,207	3,016	3,098					
GBB3-Gliederschmerzen	Neurot. Störung.	15	6,733	5,021	8,267	4,621	0,198	0,508	0,096	0,837	0,058
	Affektive Störung.	32	6,531	4,544	5,344	3,948					
	Persönlichkeitsst.	18	4,611	2,933	4,556	3,838					
	Gesamt	65	6,046	4,314	5,800	4,254					
GBB4-Herzbeschwerden	Neurot. Störung	15	6,000	5,806	3,800	3,234	0,001	0,510	0,144	0,007	0,184
	Affektive Störung	32	3,812	3,737	2,625	2,709					
	Persönlichkeitsst.	17	2,941	2,193	2,824	2,729					
	Gesamt	64	4,094	4,105	2,953	2,836					
GBB5-Beschwerdedruck	Neurot. Störung	15	25,267	15,700	25,000	11,464	0,198	0,568	0,155	0,109	0,404
	Affektive Störung	31	21,161	13,064	17,097	13,034					
	Persönlichkeitsst.	16	18,188	10,572	16,438	10,714					
	Gesamt	62	21,387	13,206	18,839	12,423					
MMPI-Hypochondriasis	Neurot. Störung	15	7,733	4,935	7,933	4,574	0,914	0,858	0,506	0,068	0,228
	Affektive Störung	31	7,581	4,863	5,903	4,437					
	Persönlichkeitsst.	17	6,647	4,122	5,588	3,809					
	Gesamt	63	7,365	4,640	6,302	4,343					
MMPI-Depression	Neurot. Störung	15	10,200	5,759	10,400	5,193	0,260	0,278	0,994	0,105	0,246
	Affektive Störung	31	11,226	5,524	9,129	5,841					
	Persönlichkeitsst.	18	10,778	3,098	9,833	3,854					
	Gesamt	64	10,859	4,966	9,625	5,150					
MMPI-Hysteria	Neurot. Störung	15	12,067	6,029	12,667	5,678	0,242	0,620	0,716	0,046	0,106
	Affektive Störung	29	12,241	6,828	9,759	5,804					
	Persönlichkeitsst.	17	12,353	4,300	10,529	4,170					
	Gesamt	61	12,230	5,931	10,689	5,412					
MMPI-Psychasthenia	Neurot. Störung	15	11,600	6,266	11,733	6,097	0,472	0,373	0,850	0,007	0,010
	Affektive Störung	31	12,581	5,726	9,129	5,542					
	Persönlichkeitsst.	17	12,000	4,637	11,000	4,077					
	Gesamt	63	12,190	5,518	10,254	5,352					
VEV	Neurot. Störung	14	207,857	36,393	202,571	66,729	0,866	0,000	0,591	0,340	0,468
	Affektive Störung	30	197,900	35,720	205,767	38,059					
	Persönlichkeitsst.	18	186,389	29,932	202,500	28,814					
	Gesamt	62	196,806	34,629	204,097	43,262					

**Tabelle 46: Varianzanalyse (Manova): Diagnosegruppen "Neurotische Störung", "Affektive Störung",  
"Spezifische Persönlichkeitsstörung", ihre Entwicklung von t1 bis t5 in der Untersuchungsgruppe, gemessen mit den  
GBB-Skalen 1 bis 5 und den MMPI(K)-Skalen HS, D, HY und PT**

			t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	Neurot. Störung	18	12,056	6,458	8,167	4,938	0,879	1,000	0,448	0,000	0,379
	Affektive Störung	34	11,088	5,468	6,000	5,099					
	Persönlichkeitsst.	20	9,800	6,296	6,800	5,126					
	Gesamt	72	10,972	5,931	6,764	5,073					
GBB2-Magenbeschwerden	Neurot. Störung	17	5,882	4,357	3,294	2,469	0,589	0,505	0,569	0,000	0,128
	Affektive Störung	32	6,156	4,746	3,062	3,574					
	Persönlichkeitsst.	19	4,000	3,801	3,263	3,634					
	Gesamt	68	5,485	4,440	3,176	3,305					
GBB3-Gliederschmerzen	Neurot. Störung	17	9,706	5,643	7,647	4,499	0,991	0,296	0,202	0,000	0,120
	Affektive Störung	33	9,515	5,432	5,242	4,016					
	Persönlichkeitsst.	20	6,900	4,898	5,350	5,354					
	Gesamt	70	8,814	5,401	5,857	4,595					
GBB4-Herzbeschwerden	Neurot. Störung	18	7,056	4,764	4,000	3,481	0,193	0,048	0,310	0,000	0,101
	Affektive Störung	33	6,333	3,680	2,485	2,539					
	Persönlichkeitsst.	19	4,684	3,575	3,579	4,451					
	Gesamt	70	6,071	3,998	3,171	3,405					
GBB5-Beschwerdedruck	Neurot. Störung	16	32,938	15,737	23,000	11,378	1,000	0,287	0,372	0,000	0,055
	Affektive Störung	30	32,533	15,529	16,633	13,032					
	Persönlichkeitsst.	18	24,778	13,910	19,056	15,569					
	Gesamt	64	30,453	15,330	18,906	13,459					
MMPI-Hypochondriasis	Neurot. Störung	16	9,563	4,305	7,188	4,861	0,908	1,000	0,668	0,000	0,113
	Affektive Störung	32	10,281	4,953	5,875	4,383					
	Persönlichkeitsst.	20	8,300	4,758	6,100	4,833					
	Gesamt	68	9,529	4,758	6,250	4,592					
MMPI-Depression	Neurot. Störung	17	14,588	5,635	9,471	5,547	0,359	0,697	0,917	0,000	0,286
	Affektive Störung	34	14,294	4,661	8,676	5,829					
	Persönlichkeitsst.	20	13,250	4,541	10,050	4,718					
	Gesamt	71	14,070	4,835	9,254	5,427					
MMPI-Hysteria	Neurot. Störung	17	15,176	5,318	11,588	6,124	0,609	0,854	0,887	0,000	0,141
	Affektive Störung	29	16,103	6,008	9,483	5,673					
	Persönlichkeitsst.	20	14,400	5,041	10,950	5,414					
	Gesamt	66	15,348	5,521	10,470	5,701					
MMPI-Psychasthenia	Neurot. Störung	17	15,294	5,807	11,000	6,062	0,785	0,540	0,776	0,000	0,042
	Affektive Störung	31	15,871	5,494	8,677	5,653					
	Persönlichkeitsst.	19	13,842	4,868	10,368	4,585					
	Gesamt	67	15,149	5,394	9,746	5,498					



**Tabelle 47: Varianzanalyse (Manova): Geschlecht der Patienten, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Einjahreszeitraum von t1 bis t3 in der Untersuchungsgruppe**

Patienten- geschlecht			t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	w	80	11,075	5,807	7,300	5,065	0,440	0,246	0,703	0,000	0,390
	m	29	10,172	5,224	7,414	5,937					
	Gesamt	109	10,835	5,648	7,330	5,283					
GBB2-Magenbeschwerden	w	79	5,360	4,429	3,418	3,157	0,940	0,880	0,675	0,000	0,459
	m	30	5,433	4,384	4,000	3,999					
	Gesamt	109	5,385	4,379	3,578	3,397					
GBB3-Gliederschmerzen	w	76	8,868	5,227	6,145	3,730	0,775	0,050	0,098	0,000	0,592
	m	30	7,100	5,492	4,900	4,908					
	Gesamt	106	8,368	5,372	5,792	4,112					
GBB4-Herzbeschwerden	w	77	5,831	3,592	3,870	3,809	0,791	0,703	0,655	0,000	0,854
	m	29	5,586	3,727	3,483	3,612					
	Gesamt	106	5,764	3,614	3,764	3,743					
GBB5-Beschwerdedruck	w	29	30,443	13,991	20,057	11,347	0,872	0,059	0,633	0,000	0,438
	m	70	28,103	14,319	19,931	14,899					
	Gesamt	99	29,758	14,055	20,020	12,413					
MMPI-Hypochondriasis	w	27	9,823	4,346	7,620	4,333	0,518	0,265	0,190	0,000	0,984
	m	79	8,630	4,757	6,444	5,046					
	Gesamt	106	9,519	4,462	7,321	4,535					
MMPI-Depression	w	79	14,354	4,194	10,734	5,004	0,543	0,507	0,994	0,000	0,222
	m	29	13,793	4,562	11,310	5,484					
	Gesamt	108	14,204	4,282	10,889	5,118					
MMPI-Hysteria	w	73	16,164	4,941	12,644	5,613	0,777	0,233	0,107	0,000	0,290
	m	27	13,778	5,147	11,481	6,664					
	Gesamt	100	15,520	5,084	12,330	5,903					
MMPI-Psychasthenia	w	78	14,936	4,974	11,718	5,573	0,592	0,673	0,701	0,000	0,390
	m	27	14,889	5,359	12,593	5,911					
	Gesamt	105	14,924	5,049	11,943	5,646					

**Tabelle 48: Varianzanalyse (Manova): Geschlecht der Patienten, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT und dem VEV im Zweijahreszeitraum von t3 bis t5 in der Untersuchungsgruppe**

	Patienten- geschlecht		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	w	45	7,200	4,708	6,956	5,099	0,063	0,784	0,484	0,149	0,322
	m	20	8,650	6,580	7,350	5,354					
	Gesamt	65	7,646	5,343	7,077	5,140					
GBB2-Magenbeschwerden	w	43	3,256	2,759	3,000	3,177	0,058	0,751	0,524	0,056	0,220
	m	20	4,200	3,955	3,050	3,000					
	Gesamt	63	3,556	3,207	3,016	3,098					
GBB3-Gliederschmerzen	w	45	6,222	3,674	6,356	3,874	0,020	0,196	0,256	0,332	0,217
	m	20	5,650	5,585	4,550	4,883					
	Gesamt	65	6,046	4,314	5,800	4,254					
GBB4-Herzbeschwerden	w	44	4,318	4,219	3,068	2,905	0,663	0,741	0,526	0,016	0,688
	m	20	3,600	3,899	2,700	2,736					
	Gesamt	64	4,094	4,105	2,953	2,836					
GBB5-Beschwerdedruck	w	42	21,048	11,436	19,405	12,081	0,044	0,587	0,915	0,019	0,270
	m	20	22,100	16,641	17,650	13,355					
	Gesamt	62	21,387	13,206	18,839	12,423					
MMPI-Hypochondriasis	w	43	7,651	4,248	6,372	4,175	0,169	0,446	0,621	0,050	0,473
	m	20	6,750	5,457	6,150	4,793					
	Gesamt	63	7,365	4,640	6,302	4,343					
MMPI-Depression	w	44	10,500	4,718	8,932	5,258	0,388	0,514	0,172	0,089	0,375
	m	20	11,650	5,518	11,150	4,671					
	Gesamt	64	10,859	4,966	9,625	5,150					
MMPI-Hysteria	w	41	12,341	5,457	10,415	5,291	0,191	0,653	0,863	0,039	0,356
	m	20	12,000	6,959	11,250	5,748					
	Gesamt	61	12,230	5,931	10,689	5,412					
MMPI-Psychasthenia	w	43	12,047	5,678	9,512	5,629	0,694	0,182	0,309	0,005	0,092
	m	20	12,500	5,287	11,850	4,416					
	Gesamt	63	12,190	5,518	10,254	5,352					
VEV	w	42	204,167	33,240	210,762	48,730	0,979	0,000	0,008	0,251	0,871
	m	20	181,350	33,080	190,100	24,087					
	Gesamt	62	196,806	34,629	204,097	43,262					

**Tabelle 49: Varianzanalyse (Manova): Geschlecht der Patienten, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Dreijahreszeitraum von t1 bis t5 in der Untersuchungsgruppe**

	Patienten- geschlecht		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	w	49	11,592	6,017	6,714	4,971	0,810	0,673	0,507	0,000	0,250
	m	24	10,000	5,779	6,708	5,336					
	Gesamt	73	11,068	5,947	6,712	5,057					
GBB2-Magenbeschwerden	w	46	5,304	4,345	3,022	3,116	0,721	0,339	0,621	0,000	0,917
	m	23	5,783	4,622	3,391	3,667					
	Gesamt	69	5,464	4,411	3,145	3,291					
GBB3-Gliederschmerzen	w	47	9,681	5,070	6,064	3,870	0,468	0,021	0,160	0,000	0,176
	m	24	7,292	5,737	5,417	5,755					
	Gesamt	71	8,873	5,385	5,845	4,563					
GBB4-Herzbeschwerden	w	47	6,362	3,703	2,830	2,721	0,264	0,005	0,946	0,000	0,121
	m	24	5,542	4,482	3,750	4,426					
	Gesamt	71	6,085	3,970	3,141	3,390					
GBB5-Beschwerdedruck	w	42	32,214	14,937	18,333	11,807	0,752	0,080	0,616	0,000	0,136
	m	23	27,739	15,801	19,565	16,169					
	Gesamt	65	30,631	15,277	18,769	13,399					
MMPI-Hypochondriasis	w	48	10,188	4,360	6,063	4,123	0,209	0,062	0,456	0,000	0,047
	m	21	8,238	5,421	6,381	5,705					
	Gesamt	69	9,594	4,754	6,159	4,620					
MMPI-Depression	w	49	14,776	4,473	8,286	5,420	0,225	0,932	0,744	0,000	0,001
	m	23	12,870	5,505	10,913	5,342					
	Gesamt	72	14,167	4,870	9,125	5,498					
MMPI-Hysteria	w	45	16,556	4,846	9,956	5,334	0,179	0,187	0,326	0,000	0,007
	m	22	13,091	6,140	11,045	6,729					
	Gesamt	67	15,418	5,508	10,313	5,800					
MMPI-Psychasthenia	w	47	15,426	5,484	8,809	5,682	0,965	0,416	0,383	0,000	0,045
	m	21	14,952	5,527	11,429	4,925					
	Gesamt	68	15,279	5,460	9,618	5,558					

**Tabelle 50: Varianzanalyse (Manova): Alter der Patienten zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Alter		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	bis 30	32	10,625	5,369	7,406	5,535	0,051	1,000	0,971	0,000	0,500
	31-40	59	10,797	5,346	7,559	5,210					
	über 40	18	11,333	7,227	6,444	5,272					
	Gesamt	109	10,835	5,648	7,330	5,283					
GBB2-Magenbeschwerden	bis 30	32	5,844	4,739	4,375	3,957	0,721	0,090	0,471	0,000	0,781
	31-40	60	5,267	4,286	3,333	3,182					
	über 40	17	4,941	4,308	2,941	2,883					
	Gesamt	109	5,385	4,397	3,578	3,397					
GBB3-Gliederschmerzen	bis 30	32	7,563	4,529	5,875	4,179	0,015	1,000	0,028	0,000	0,176
	31-40	56	7,804	4,989	5,232	4,036					
	über 40	18	11,556	6,879	7,389	4,017					
	Gesamt	106	8,368	5,372	5,792	4,112					
GBB4-Herzbeschwerden	bis 30	31	5,935	3,915	4,226	4,105	0,715	0,050	0,362	0,000	0,667
	31-40	59	5,864	3,496	3,898	3,894					
	über 40	16	5,063	3,586	2,375	1,784					
	Gesamt	106	5,764	3,614	3,764	3,743					
GBB5-Beschwerdedruck	bis 30	31	29,645	12,935	21,355	13,438	0,069	0,488	0,811	0,000	0,296
	31-40	52	28,788	13,554	19,577	12,401					
	über 40	16	33,125	17,780	18,875	10,813					
	Gesamt	99	29,758	14,055	20,020	12,413					
MMPI-Hypochondriasis	bis 30	30	10,400	4,056	7,900	5,175	0,079	0,165	0,497	0,000	0,191
	31-40	60	9,017	4,343	7,350	4,372					
	über 40	16	9,750	5,544	6,125	3,845					
	Gesamt	106	9,519	4,462	7,321	4,535					
MMPI-Depression	bis 30	31	15,032	3,371	11,484	5,259	0,007	0,910	0,263	0,000	0,583
	31-40	60	14,050	4,244	11,083	4,921					
	über 40	17	13,235	5,696	9,118	5,464					
	Gesamt	108	14,204	4,282	10,889	5,118					
MMPI-Hysteria	bis 30	30	16,867	4,939	13,300	7,028	0,206	0,067	0,191	0,000	0,311
	31-40	55	15,000	4,823	12,436	5,343					
	über 40	15	14,733	6,076	10,000	5,113					
	Gesamt	100	15,520	5,084	12,330	5,903					
MMPI-Psychasthenia	bis 30	31	15,226	5,064	12,258	6,099	0,606	0,428	0,772	0,000	0,409
	31-40	57	14,772	4,946	12,193	5,125					
	über 40	17	14,882	5,633	10,529	6,559					
	Gesamt	105	14,924	5,049	11,943	5,646					

**Tabelle 51: Varianzanalyse (Manova): Alter der Patienten zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT und dem VEV im Zweijahreszeitraum von t3 bis t5 in der Untersuchungsgruppe**

	Alter		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	bis 30	15	7,600	5,488	7,667	5,150	1,000	0,295	0,899	0,288	0,692
	31-40	38	7,737	5,371	7,132	4,839					
	über 40	12	7,417	5,534	6,167	6,308					
	Gesamt	65	7,646	5,343	7,077	5,140					
GBB2-Magenbeschwerden	bis 30	15	4,333	3,773	3,800	3,610	0,373	0,392	0,462	0,046	0,378
	31-40	38	3,105	2,994	2,842	2,955					
	über 40	10	4,100	3,107	2,500	2,915					
	Gesamt	63	3,556	3,207	3,016	3,098					
GBB3-Gliederschmerzen	bis 30	15	5,667	4,530	6,133	4,912	0,591	0,502	0,359	0,292	0,129
	31-40	38	5,447	4,329	5,526	4,105					
	über 40	12	8,417	3,397	6,250	4,159					
	Gesamt	65	6,046	4,314	5,800	4,254					
GBB4-Herzbeschwerden	bis 30	15	4,267	4,728	3,067	2,549	0,066	0,186	0,351	0,021	0,967
	31-40	38	4,421	4,354	3,237	3,097					
	über 40	11	2,727	1,489	1,818	2,089					
	Gesamt	64	4,094	4,105	2,953	2,836					
GBB5-Beschwerdedruck	bis 30	15	21,867	15,556	20,667	14,435	0,188	0,604	0,914	0,016	0,258
	31-40	38	20,711	13,127	18,737	11,638					
	über 40	9	23,444	10,101	16,222	13,103					
	Gesamt	62	21,387	13,206	18,839	12,423					
MMPI-Hypochondriasis	bis 30	15	7,533	5,579	7,800	5,185	0,219	0,278	0,618	0,062	0,233
	31-40	37	7,432	4,525	6,000	3,993					
	über 40	11	6,909	3,986	5,273	4,125					
	Gesamt	63	7,365	4,640	6,302	4,343					
MMPI-Depression	bis 30	15	10,200	5,335	10,800	5,321	1,000	0,688	0,639	0,080	0,172
	31-40	38	11,316	4,839	9,658	4,878					
	über 40	11	10,182	5,193	7,909	5,839					
	Gesamt	64	10,859	4,966	9,625	5,150					
MMPI-Hysteria	bis 30	14	12,643	7,323	13,786	5,632	0,184	0,849	0,196	0,096	0,044
	31-40	37	12,568	5,565	10,162	5,019					
	über 40	10	10,400	5,358	8,300	5,122					
	Gesamt	61	12,230	5,931	10,689	5,412					
MMPI-Psychasthenia	bis 30	15	11,600	6,057	11,000	4,645	0,981	0,024	0,873	0,001	0,236
	31-40	37	12,432	5,247	10,378	4,946					
	über 40	11	12,182	6,129	8,818	7,494					
	Gesamt	63	12,190	5,518	10,254	5,352					
VEV	bis 30	14	203,071	29,042	210,429	34,435	0,446	0,260	0,341	0,173	0,698
	31-40	36	193,639	37,548	197,444	47,232					
	über 40	12	199,000	32,850	216,667	38,874					
	Gesamt	62	196,806	34,629	204,097	43,262					

**Tabelle 52: Varianzanalyse (Manova): Alter der Patienten zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Alter		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	bis 30	16	10,563	6,460	7,375	5,110	0,391	0,284	0,509	0,000	0,693
	31-40	42	11,548	5,325	7,000	4,701					
	über 40	14	9,714	7,184	5,357	6,160					
	Gesamt	72	10,972	5,931	6,764	5,073					
GBB2-Magenbeschwerden	bis 30	16	6,375	4,856	3,750	3,493	0,825	0,845	0,609	0,000	0,575
	31-40	40	5,025	4,317	3,125	3,376					
	über 40	12	5,833	4,448	2,583	2,937					
	Gesamt	68	5,485	4,440	3,176	3,305					
GBB3-Gliederschmerzen	bis 30	16	7,750	5,040	6,313	4,799	0,547	0,874	0,515	0,000	0,122
	31-40	40	8,425	5,218	5,650	4,817					
	über 40	14	11,143	6,011	5,929	3,931					
	Gesamt	70	8,814	5,401	5,857	4,595					
GBB4-Herzbeschwerden	bis 30	16	5,750	3,768	3,563	3,162	0,106	0,022	0,731	0,000	0,390
	31-40	41	6,146	3,759	3,439	3,847					
	über 40	13	6,231	5,183	1,846	1,573					
	Gesamt	70	6,071	3,998	3,171	3,405					
GBB5-Beschwerdedruck	bis 30	16	30,438	15,904	21,000	14,010	0,166	1,000	0,930	0,000	0,405
	31-40	37	29,757	13,859	18,865	13,710					
	über 40	11	32,818	20,044	16,000	12,402					
	Gesamt	64	30,453	15,330	18,906	13,459					
MMPI-Hypochondriasis	bis 30	14	10,571	4,863	7,929	5,356	0,121	0,213	0,331	0,000	0,356
	31-40	41	9,244	4,259	6,220	4,441					
	über 40	13	9,308	6,223	4,538	3,799					
	Gesamt	68	9,529	4,758	6,250	4,592					
MMPI-Depression	bis 30	15	15,067	4,044	10,800	5,321	0,003	1,000	0,180	0,000	0,626
	31-40	42	14,071	4,239	9,476	5,320					
	über 40	14	13,000	7,016	6,929	5,484					
	Gesamt	71	14,070	4,835	9,254	5,427					
MMPI-Hysteria	bis 30	13	16,846	6,012	13,692	5,851	0,113	0,689	0,051	0,000	0,264
	31-40	39	15,154	4,586	10,487	5,486					
	über 40	14	14,500	7,377	7,429	4,751					
	Gesamt	66	15,348	5,521	10,470	5,701					
MMPI-Psychasthenia	bis 30	14	16,286	5,622	10,929	4,811	1,000	0,136	0,474	0,000	0,600
	31-40	39	14,872	5,287	9,923	5,208					
	über 40	14	14,786	5,713	8,071	6,810					
	Gesamt	67	15,149	5,394	9,746	5,498					

**Tabelle 53: Varianzanalyse (Manova): Dyaden Therapeut/Patient (T:m/P:m,T:w/P:m,T:w/P:w/T:m/P:w),  
Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT  
im Einjahreszeitraum von t1 bis t3 in der Untersuchungsgruppe**

	Dyade Therapeut/ Patient		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	M/M	15	9,467	4,291	6,800	4,799	0,585	0,023	0,087	0,000	0,863
	W/M	14	10,629	6,145	8,071	7,087					
	W/W	44	9,909	5,894	6,159	5,287					
	M/W	36	12,500	5,443	8,694	4,465					
	Gesamt										
GBB2-Magenbeschwerden	M/M	16	5,562	4,802	3,188	3,582	0,819	0,131	0,406	0,000	0,120
	W/M	14	5,286	4,027	4,929	4,358					
	W/W	45	4,578	4,377	3,111	3,142					
	M/W	34	6,412	4,342	3,824	3,176					
	Gesamt										
GBB3-Gliederschmerzen	M/M	16	5,875	3,964	2,750	2,206	0,073	0,000	0,035	0,000	0,377
	W/M	14	8,500	6,723	7,357	5,995					
	W/W	44	8,409	5,050	6,205	4,101					
	M/W	32	9,500	5,594	6,063	3,212					
	Gesamt										
GBB4-Herzbeschwerden	M/M	15	4,600	2,898	2,400	2,354	0,237	0,173	0,170	0,000	0,728
	W/M	14	6,643	4,308	4,643	4,396					
	W/W	44	5,227	2,972	3,659	3,888					
	M/W	33	6,636	4,197	4,152	3,743					
	Gesamt										
GBB5-Beschwerdedruck	M/M	15	25,067	11,361	15,200	10,241	0,239	0,004	0,065	0,000	0,603
	W/M	14	31,357	16,750	25,000	17,665					
	W/W	42	27,810	12,748	18,500	11,327					
	M/W	28	34,393	15,051	22,393	11,170					
	Gesamt										
MMPI-Hypochondriasis	M/M	14	7,500	3,716	5,643	4,378	0,051	0,111	0,018	0,000	0,378
	W/M	13	9,846	5,565	7,308	5,765					
	W/W	44	8,455	3,891	6,932	4,212					
	M/W	35	11,543	4,327	8,486	4,388					
	Gesamt										
MMPI-Depression	M/M	16	13,457	3,366	10,563	3,829	0,009	0,007	0,127	0,000	0,606
	W/M	13	14,231	5,833	12,231	7,085					
	W/W	43	13,326	4,455	9,814	4,832					
	M/W	36	15,583	3,541	11,833	5,051					
	Gesamt										
MMPI-Hysteria	M/M	15	13,533	4,688	11,400	5,235	0,297	0,011	0,032	0,000	0,623
	W/M	12	14,083	5,869	11,583	8,372					
	W/W	39	14,641	4,793	11,564	5,428					
	M/W	34	17,912	4,575	13,882	5,645					
	Gesamt										
MMPI-Psychasthenia	M/M	14	15,286	4,531	12,214	4,823	0,108	0,122	0,109	0,000	0,605
	W/M	13	14,462	6,293	13,000	7,083					
	W/W	44	13,630	4,842	10,705	5,037					
	M/W	34	16,618	4,691	13,029	6,023					
	Gesamt										

**Tabelle 54: Varianzanalyse (Manova): Dyaden Therapeut/Patient (T:m/P:m,T:w/P:m,T:w/P:w/T:m/P:w),  
Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT und dem VEV  
im Zweijahreszeitraum von t3 bis t5 in der Untersuchungsgruppe**

	Dyade Therapeut/ Patient		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	M/M	13	7,077	5,107	5,923	4,071	0,006	0,118	0,043	0,153	0,797
	W/M	7	11,571	8,344	10,000	6,708					
	W/W	25	5,920	4,554	5,680	5,218					
	M/W	20	8,800	4,503	8,550	4,582					
	Gesamt	65	7,646	5,343	7,077	5,140					
GBB2-Magenbeschwerden	M/M	13	3,462	3,950	3,077	3,252	0,471	0,137	0,477	0,012	0,021
	W/M	7	5,571	3,867	3,000	2,708					
	W/W	24	2,375	2,123	2,917	3,878					
	M/W	19	4,368	3,183	3,105	2,079					
	Gesamt	63	3,556	3,207	3,016	3,098					
GBB3-Gliederschmerzen	M/M	13	2,923	2,326	3,154	2,734	0,001	0,001	0,006	0,141	0,062
	W/M	7	10,714	6,473	7,143	6,962					
	W/W	25	6,400	4,223	6,080	4,132					
	M/W	20	6,000	2,938	6,700	3,600					
	Gesamt	65	6,046	4,314	5,800	4,254					
GBB4-Herzbeschwerden	M/M	13	2,692	2,394	2,462	2,537	0,072	0,691	0,637	0,010	0,357
	W/M	7	5,286	5,619	3,143	3,237					
	W/W	25	4,440	4,331	2,680	2,529					
	M/W	19	4,158	4,180	3,579	3,339					
	Gesamt	64	4,094	4,105	2,953	2,836					
GBB5-Beschwerdedruck	M/M	13	16,154	10,691	14,615	10,145	0,002	0,047	0,083	0,006	0,176
	W/M	7	33,143	20,724	23,286	17,375					
	W/W	24	19,375	12,268	17,417	13,590					
	M/W	18	23,278	10,127	22,056	9,440					
	Gesamt	62	21,387	13,206	18,859	12,423					
MMPI-Hypochondriasis	M/M	13	5,846	4,451	5,846	4,705	0,018	0,546	0,437	0,033	0,660
	W/M	7	8,429	7,044	6,714	5,282					
	W/W	24	6,917	4,624	5,625	4,744					
	M/W	19	8,579	3,626	7,316	3,198					
	Gesamt	63	7,365	4,640	6,302	4,343					
MMPI-Depression	M/M	13	10,692	3,881	10,769	3,140	0,005	0,070	0,190	0,059	0,444
	W/M	7	13,429	7,786	11,837	6,962					
	W/W	24	9,250	4,821	8,375	5,948					
	M/W	20	12,000	4,230	9,600	4,346					
	Gesamt	64	10,859	4,966	9,625	5,150					
MMPI-Hysteria	M/M	13	11,846	5,444	11,077	5,454	0,006	0,474	0,485	0,047	0,782
	W/M	7	12,286	9,676	11,571	6,705					
	W/W	22	11,000	5,855	9,409	5,561					
	M/W	19	13,895	4,630	11,579	4,846					
	Gesamt	61	12,230	5,931	10,689	5,412					
MMPI-Psychasthenia	M/M	13	11,462	5,093	11,692	3,199	1,000	0,268	0,072	0,003	0,206
	W/M	7	14,429	5,473	12,143	6,414					
	W/W	24	10,542	5,200	7,875	5,059					
	M/W	19	13,947	5,816	11,579	5,757					
	Gesamt	63	12,190	5,518	10,254	5,352					
VEV	M/M	13	190,385	18,455	192,000	25,299	0,030	0,000	0,031	0,137	0,389
	W/M	7	164,571	47,742	186,571	23,237					
	W/W	25	205,520	31,618	202,920	57,770					
	M/W	17	202,176	36,396	222,294	29,008					
	Gesamt	62	196,806	34,629	204,097	43,262					



**Tabelle 55: Varianzanalyse (Manova): Dyaden Therapeut/Patient (T:m/P:m,T:w/P:m,T:w/P:w/T:m/P:w), Zusammenhang mit den GBB-Skalen 1 bis 5 und den MMPI(K)-Skalen HS, D, HY, PT im Dreijahreszeitraum t1 bis t5 in der Untersuchungsgruppe**

	Dyade Therapeut/ Patient		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	M/M	16	9,187	4,736	5,500	4,258	0,164	0,083	0,065	0,000	0,671
	W/M	8	11,625	7,558	9,125	6,686					
	W/W	28	10,321	6,195	5,464	4,992					
	M/W	21	13,286	5,460	8,381	4,533					
	Gesamt	73	11,086	5,947	6,712	5,057					
GBB2-Magenbeschwerden	M/M	15	5,600	5,221	3,600	4,239	0,315	0,073	0,302	0,000	0,245
	W/M	8	6,125	3,523	3,000	2,507					
	W/W	26	4,000	3,441	2,731	3,779					
	M/W	20	7,000	4,877	3,400	1,984					
	Gesamt	69	5,464	4,411	3,145	3,291					
GBB3-Gliederschmerzen	M/M	16	6,312	4,813	4,438	5,304	0,096	0,067	0,213	0,000	0,469
	W/M	8	9,250	7,206	7,375	6,479					
	W/W	28	9,893	5,123	5,750	4,052					
	M/W	19	9,368	5,112	6,526	3,642					
	Gesamt	71	8,873	5,385	5,845	4,563					
GBB4-Herzbeschwerden	M/M	16	5,750	4,919	3,563	4,704	0,295	0,101	0,768	0,000	0,350
	W/M	8	5,125	3,720	4,125	4,086					
	W/W	28	5,786	3,131	2,643	2,468					
	M/W	19	7,211	4,366	3,105	3,107					
	Gesamt	71	6,085	3,970	3,141	3,390					
GBB5-Beschwerdedruck	M/M	15	25,400	13,292	17,400	16,326	0,201	0,491	0,255	0,000	0,430
	W/M	8	32,125	19,939	23,625	16,115					
	W/W	26	29,231	13,204	16,692	13,296					
	M/W	16	37,062	16,695	21,000	8,610					
	Gesamt	65	30,631	15,277	18,769	13,399					
MMPI-Hypochondriasis	M/M	15	8,133	4,749	6,333	5,888	0,012	0,404	0,236	0,000	0,230
	W/M	6	8,500	7,369	6,500	5,753					
	W/W	27	9,000	3,922	5,272	4,644					
	M/W	21	11,714	4,507	7,143	3,119					
	Gesamt	69	9,594	4,754	6,159	4,620					
MMPI-Depression	M/M	16	12,437	4,560	10,500	4,676	0,004	0,162	0,243	0,000	0,017
	W/M	7	13,857	7,581	11,857	6,962					
	W/W	28	13,679	5,026	7,357	6,063					
	M/W	21	16,238	3,161	9,524	4,250					
	Gesamt	72	14,167	4,870	9,125	9,125					
MMPI-Hysteria	M/M	15	13,200	5,979	10,800	6,961	0,184	0,613	0,131	0,000	0,053
	W/M	7	12,857	6,962	11,571	6,705					
	W/W	24	14,958	5,017	8,792	5,664					
	M/W	21	18,381	4,018	11,286	4,713					
	Gesamt	67	15,418	5,508	10,313	5,800					
MMPI-Psychasthenia	M/M	14	14,786	5,010	11,071	4,233	0,233	0,429	0,038	0,000	0,260
	W/M	7	15,286	6,873	12,143	6,414					
	W/W	27	13,889	5,611	7,185	5,168					
	M/W	20	17,500	4,674	11,000	5,731					
	Gesamt	68	15,279	5,460	9,618	5,558					

**Tabelle 57: Varianzanalyse (Manova): Sitzungsfrequenz zu t3 (Trennung über dem Median in niedriger und höher), Zusammenhang mit den GBB-Skalen 1 bis 5 und den Skalen des MMPI(K) HS, D, HY, PT im Einjahreszeitraum von t1 bis t3 in der Untersuchungsgruppe**

	Frequenz zu t3	N	t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
			M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	Niedrig	54	10,667	5,610	7,333	4,794	0,960	0,164	0,770	0,000	0,659
	Hoch	54	11,167	5,649	7,370	5,812					
	Gesamt	108	10,917	5,609	7,352	5,303					
GBB2-Magenbeschwerden	Niedrig	54	5,185	4,176	3,389	3,389	0,436	0,898	0,531	0,000	0,930
	Hoch	54	5,648	4,651	3,796	3,450					
	Gesamt	108	5,417	4,405	3,593	3,410					
GBB3-Gliederschmerzen	Niedrig	53	8,321	5,064	5,792	3,840	0,373	0,218	0,898	0,000	0,871
	Hoch	52	8,500	5,734	5,827	4,440					
	Gesamt	105	8,410	5,381	5,810	4,128					
GBB4-Herzbeschwerden	Niedrig	52	5,942	3,928	4,192	4,010	0,231	0,302	0,387	0,000	0,477
	Hoch	53	5,642	3,323	3,396	3,472					
	Gesamt	105	5,790	3,621	3,790	3,751					
GBB5-Beschwerdedruck	Niedrig	47	29,064	12,470	20,404	11,409	0,154	0,259	0,812	0,000	0,384
	Hoch	51	30,765	15,350	19,824	13,446					
	Gesamt	98	29,940	13,997	20,102	12,450					
MMPI-Hypochondriasis	Niedrig	54	10,074	3,885	7,593	4,428	0,071	0,654	0,328	0,000	0,535
	Hoch	51	9,059	4,966	7,059	4,713					
	Gesamt	105	9,581	4,437	7,333	4,555					
MMPI-Depression	Niedrig	55	14,291	4,237	10,855	5,108	0,972	0,874	0,949	0,000	0,894
	Hoch	52	14,288	4,216	10,962	5,220					
	Gesamt	107	14,290	4,207	10,907	5,139					
MMPI-Hysteria	Niedrig	52	15,923	4,410	12,481	5,648	0,074	0,499	0,658	0,000	0,627
	Hoch	47	15,234	5,707	12,298	6,224					
	Gesamt	99	15,596	5,053	12,394	5,898					
MMPI-Psychasthenia	Niedrig	53	14,774	5,380	12,057	6,071	0,359	0,316	0,904	0,000	0,542
	Hoch	51	15,176	4,727	11,882	5,271					
	Gesamt	104	14,971	5,050	11,971	5,666					

**Tabelle 58: Varianzanalyse (Manova): Sitzungsfrequenz zu t4 (Trennung über dem Median in niedriger und höher), Zusammenhang mit den GBB-Skalen 1 bis 5, den Skalen des MMPI(K) HS, D, HY, PT und dem VEV im Einjahreszeitraum von t3 bis t4 in der Untersuchungsgruppe**

	Frequenz zu t4		t3		t4		t3	t4	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	Niedrig	40	6,600	5,256	7,075	5,663	0,780	0,109	0,847	0,257	0,039
	Hoch	38	7,842	5,022	6,237	4,352					
	Gesamt	78	7,205	5,148	6,667	5,052					
GBB2-Magenbeschwerden	Niedrig	38	2,737	2,789	2,789	3,330	0,364	0,178	0,804	0,270	0,205
	Hoch	37	3,297	2,541	2,541	2,662					
	Gesamt	75	3,042	2,667	2,667	3,002					
GBB3-Gliederschmerzen	Niedrig	40	5,550	4,761	6,525	5,129	0,105	0,040	0,370	0,885	0,002
	Hoch	37	5,662	3,647	4,730	3,656					
	Gesamt	77	5,584	4,234	5,662	4,544					
GBB4-Herzbeschwerden	Niedrig	39	3,359	3,808	3,897	3,789	0,958	0,166	0,814	0,639	0,031
	Hoch	36	3,861	3,841	3,028	3,009					
	Gesamt	75	3,600	3,806	3,480	3,442					
GBB5-Beschwerdedruck	Niedrig	38	17,421	11,533	19,895	14,080	0,894	0,045	0,930	0,412	0,002
	Hoch	36	20,500	11,795	16,361	10,023					
	Gesamt	74	18,919	11,684	18,176	12,321					
MMPI-Hypochondriasis	Niedrig	38	7,579	4,780	7,158	4,722	0,510	0,100	0,187	0,003	0,064
	Hoch	39	7,000	4,292	5,256	3,603					
	Gesamt	77	7,286	4,519	6,195	4,274					
MMPI-Depression	Niedrig	37	10,838	5,408	10,459	5,178	0,549	0,238	0,775	0,460	0,805
	Hoch	37	10,432	4,891	10,243	4,245					
	Gesamt	74	10,635	5,125	10,351	4,703					
MMPI-Hysteria	Niedrig	37	12,378	6,193	11,649	5,544	0,563	0,439	0,450	0,008	0,248
	Hoch	33	11,970	5,604	10,152	4,848					
	Gesamt	70	12,186	5,884	10,943	5,244					
MMPI-Psychasthenia	Niedrig	35	11,771	6,073	11,000	6,217	0,645	0,272	0,945	0,155	0,787
	Hoch	38	11,737	5,622	11,211	5,168					
	Gesamt	73	11,753	5,802	11,110	5,656					
VEV	Niedrig	37	203,514	37,591	204,541	33,193	0,885	0,849	0,498	0,009	0,020
	Hoch	39	190,205	36,698	207,897	32,861					
	Gesamt	76	196,684	37,490	206,263	33,197					

**Tabelle 59: Varianzanalyse (Manova): Sitzungsfrequenz zu t5 (Trennung über dem Median in niedriger und höher), Zusammenhang mit den GBB-Skalen 1 bis 5, den Skalen des MMPI(K) HS, D, HY, PT und dem VEV im Einjahreszeitraum von t4 bis t5 in der Untersuchungsgruppe**

	Frequenz zu t5		t4		t5		t4	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	Niedrig	34	7,029	5,507	6,529	4,587	0,728	0,394	0,822	0,892	0,378
	Hoch	30	6,867	5,171	7,233	5,354					
	Gesamt	64	6,953	5,311	6,859	4,934					
GBB2-Magenbeschwerden	Niedrig	33	3,030	3,340	3,333	3,789	0,623	0,134	0,913	0,464	0,126
	Hoch	27	3,519	3,662	2,667	2,855					
	Gesamt	60	3,250	3,467	3,033	3,390					
GBB3-Gliederschmerzen	Niedrig	34	6,235	5,052	5,500	4,888	0,581	0,185	0,708	0,899	0,135
	Hoch	29	5,966	4,571	6,586	3,841					
	Gesamt	63	6,111	4,800	6,000	4,436					
GBB4-Herzbeschwerden	Niedrig	34	4,088	3,808	2,971	3,746	0,283	0,160	0,875	0,042	0,288
	Hoch	28	3,571	3,841	3,214	2,898					
	Gesamt	62	3,855	3,806	3,081	3,365					
GBB5-Beschwerdedruck	Niedrig	33	20,212	11,533	18,515	13,677	0,304	0,438	0,980	0,475	0,395
	Hoch	27	19,370	11,795	19,519	11,827					
	Gesamt	60	19,833	11,684	18,967	12,781					
MMPI-Hypochondriasis	Niedrig	34	6,500	4,780	5,912	4,870	0,515	0,414	0,808	0,575	0,482
	Hoch	30	6,433	4,292	6,500	4,200					
	Gesamt	64	6,469	4,519	6,188	4,542					
MMPI-Depression	Niedrig	32	9,438	5,408	8,937	5,512	0,153	0,659	0,207	0,051	0,331
	Hoch	30	11,433	4,891	9,967	5,082					
	Gesamt	62	10,403	5,125	9,435	5,290					
MMPI-Hysteria	Niedrig	32	10,469	6,193	9,719	5,985	0,328	0,452	0,230	0,115	0,779
	Hoch	28	12,214	5,604	11,143	5,198					
	Gesamt	60	11,283	5,884	10,383	5,630					
MMPI-Psychasthenia	Niedrig	33	10,667	6,073	9,485	5,585	0,718	0,823	0,252	0,004	0,471
	Hoch	29	12,586	5,622	10,655	5,360					
	Gesamt	62	11,565	5,802	10,032	5,468					
VEV	Niedrig	33	215,152	34,717	206,970	52,110	0,406	0,030	0,319	0,467	0,425
	Hoch	31	202,645	29,855	203,032	35,019					
	Gesamt	64	209,094	32,807	205,062	44,353					

**Tabelle 60: Varianzanalyse (Manova): Sitzungsfrequenz zu t5 (Trennung über dem Median in niedriger und höher), Zusammenhang mit den GBB-Skalen 1 bis 5, den Skalen des MMPI(K) HS, D, HY, PT und dem VEV im Zweijahreszeitraum von t3 bis t5 in der Untersuchungsgruppe**

	Frequenz zu t5		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH. P	COCH. P	F	F	F
GBB1-Erschöpfungsneigung	Niedrig	34	6,941	4,566	6,765	4,684	0,113	0,292	0,381	0,234	0,403
	Hoch	31	8,419	6,065	7,419	5,656					
	Gesamt	65	7,646	5,343	7,077	5,140					
GBB2-Magenbeschwerden	Niedrig	33	3,061	2,692	3,121	3,267	0,092	0,576	0,574	0,090	0,061
	Hoch	30	4,100	3,661	2,900	2,952					
	Gesamt	63	3,550	3,207	3,016	3,098					
GBB3-Gliederschmerzen	Niedrig	34	5,265	3,792	5,029	4,338	0,214	0,713	0,090	0,596	0,980
	Hoch	31	6,903	4,735	6,645	4,062					
	Gesamt	65	6,046	4,314	5,800	4,254					
GBB4-Herzbeschwerden	Niedrig	34	3,794	3,983	2,588	2,595	0,691	0,346	0,373	0,006	0,863
	Hoch	30	4,433	4,281	3,367	3,079					
	Gesamt	64	4,094	4,105	2,953	2,836					
GBB5-Beschwerdedruck	Niedrig	33	19,030	11,351	17,455	11,659	0,153	0,483	0,191	0,031	0,384
	Hoch	29	24,069	14,789	20,414	13,268					
	Gesamt	62	21,387	13,206	18,839	12,423					
MMPI-Hypochondriasis	Niedrig	33	7,030	4,268	5,939	4,358	0,346	0,994	0,489	0,019	0,948
	Hoch	30	7,733	5,065	6,700	4,364					
	Gesamt	63	7,365	4,640	6,302	4,343					
MMPI-Depression	Niedrig	34	10,118	5,014	8,588	5,118	0,861	0,909	0,096	0,033	0,574
	Hoch	30	11,770	4,858	10,800	5,013					
	Gesamt	64	10,859	4,966	9,625	5,150					
MMPI-Hysteria	Niedrig	34	11,265	5,675	9,706	5,272	0,676	0,874	0,100	0,013	0,973
	Hoch	27	13,444	6,129	11,926	5,427					
	Gesamt	61	12,230	5,931	10,689	5,412					
MMPI-Psychasthenia	Niedrig	33	11,545	6,235	9,424	5,477	0,096	0,731	0,225	0,001	0,713
	Hoch	30	12,900	4,604	11,167	5,147					
	Gesamt	63	12,190	5,518	10,254	5,352					
VEV	Niedrig	31	207,968	30,724	205,516	50,468	0,464	0,057	0,109	0,234	0,113
	Hoch	31	185,645	35,169	202,677	35,418					
	Gesamt	62	196,806	34,629	204,097	43,262					

**Tabelle 61a: Varianzanalyse (Manova): Therapeutengeschlecht, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Therapeutengeschlecht		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
<b>GBB1-Erschöpfungsneigung</b>	Weiblich	58	10,155	5,917	6,621	5,764					
	Männlich	51	11,608	5,277	8,137	4,600					
	Gesamt	109	10,835	5,648	7,330	5,283	0,403	0,100	0,104	0,000	0,951
<b>GBB2-Magenbeschwerden</b>	Weiblich	59	4,746	4,273	3,542	3,515					
	Männlich	50	6,140	4,463	3,620	3,288					
	Gesamt	109	5,385	4,397	3,578	3,397	0,751	0,626	0,287	0,000	0,033
<b>GBB3-Gliederschmerzen</b>	Weiblich	58	8,431	5,436	6,483	4,597					
	Männlich	48	8,292	5,351	4,958	3,294					
	Gesamt	106	8,368	5,372	5,792	4,112	0,911	0,018	0,315	0,000	0,116
<b>GBB4-Herzbeschwerden</b>	Weiblich	58	5,569	3,357	3,897	3,999					
	Männlich	48	6,000	3,925	3,604	3,444					
	Gesamt	106	5,764	3,614	3,764	3,743	0,263	0,285	0,913	0,000	0,296
<b>GBB5-Beschwerdedruck</b>	Weiblich	56	28,696	13,779	20,125	13,319					
	Männlich	43	31,140	14,453	19,884	11,278					
	Gesamt	99	29,758	14,055	20,020	12,413	0,739	0,248	0,642	0,000	0,305
<b>MMPI-Hypochondriasis</b>	Weiblich	57	8,772	4,314	7,018	4,557					
	Männlich	49	10,388	4,518	7,673	4,529					
	Gesamt	106	9,519	4,462	7,321	4,535	0,740	0,965	0,150	0,000	0,216
<b>MMPI-Depression</b>	Weiblich	56	13,536	4,767	10,375	5,462					
	Männlich	52	14,923	3,597	11,442	4,709					
	Gesamt	108	14,204	4,282	10,889	5,118	0,043	0,283	0,130	0,000	0,699
<b>MMPI-Hysteria</b>	Weiblich	51	14,510	5,009	11,569	6,149					
	Männlich	49	16,571	4,996	13,122	5,589					
	Gesamt	100	15,520	5,084	12,330	5,903	0,985	0,506	0,064	0,000	0,622
<b>MMPI-Psychasthenia</b>	Weiblich	57	13,825	5,159	11,228	5,584					
	Männlich	48	16,229	4,637	12,792	5,660					
	Gesamt	105	14,924	5,049	11,943	5,646	0,444	0,922	0,034	0,000	0,371

**Tabelle 61b: Varianzanalyse (Manova): Therapeutengeschlecht, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

			t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
	Therapeutengeschlecht										
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Weiblich	51	3,686	1,954	3,980	1,913	0,427	0,507	0,369	0,128	0,806
	Männlich	47	3,426	1,741	3,638	1,737					
	Gesamt	98	3,561	1,850	3,816	1,830					
MMPI-Validity	Weiblich	48	4,333	3,238	4,125	2,863	0,030	0,065	0,085	0,008	0,039
	Männlich	43	6,233	4,493	4,581	3,781					
	Gesamt	91	5,231	3,975	4,341	3,317					
MMPI-Correction	Weiblich	55	8,927	3,563	9,891	3,710	0,279	0,815	0,382	0,000	0,502
	Männlich	51	8,176	3,064	9,549	3,833					
	Gesamt	106	8,566	3,338	9,726	3,756					
MMPI-Psychopathic Deviate	Weiblich	55	11,727	4,390	9,891	4,541	0,571	0,808	0,013	0,000	0,676
	Männlich	44	13,886	4,047	11,727	4,385					
	Gesamt	99	12,687	4,356	10,707	4,543					
MMPI-Masculinity/Femininity	Weiblich	57	14,333	2,918	14,491	3,208	0,643	0,856	0,052	0,015	0,001
	Männlich	46	16,109	3,114	15,000	3,127					
	Gesamt	103	15,126	3,121	14,718	3,167					
MMPI-Paranoia	Weiblich	51	7,078	3,117	6,510	3,379	0,028	0,182	0,107	0,001	0,143
	Männlich	47	8,660	4,295	7,170	4,104					
	Gesamt	98	7,837	3,793	6,827	3,739					
MMPI-Schizophrenia	Weiblich	51	9,000	4,427	6,902	4,912	0,333	0,751	0,135	0,000	0,305
	Männlich	46	10,761	5,095	7,761	4,691					
	Gesamt	97	9,835	4,812	7,309	4,803					
MMPI-Hypomania	Weiblich	54	9,852	2,987	9,463	3,898	0,761	0,377	0,110	0,071	0,669
	Männlich	43	10,977	2,858	10,349	3,429					
	Gesamt	97	10,351	2,969	9,856	3,705					
MMPI-Social Introversion	Weiblich	57	13,860	4,557	13,053	4,879	0,917	0,640	0,871	0,003	0,381
	Männlich	51	14,314	4,492	12,863	4,574					
	Gesamt	108	14,074	4,511	12,963	4,716					

**Tabelle 62a: Varianzanalyse (Manova): Therapeutengeschlecht, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Therapeutengeschlecht		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
GBB1-Erschöpfungsneigung	Weiblich	32	7,156	5,930	6,625	5,752					
	Männlich	33	8,121	4,748	7,515	4,515					
	Gesamt	65	7,646	5,343	7,077	5,140	0,214	0,176	0,446	0,252	0,940
GBB2-Magenbeschwerden	Weiblich	31	3,097	2,879	2,935	3,605					
	Männlich	32	4,000	3,483	3,094	2,570					
	Gesamt	63	3,556	3,207	3,016	3,098	0,295	0,064	0,465	0,118	0,273
GBB3-Gliederschmerzen	Weiblich	32	7,344	5,020	6,313	4,775					
	Männlich	33	4,788	3,080	5,303	3,687					
	Gesamt	65	6,046	4,314	5,800	4,254	0,007	0,149	0,063	0,570	0,092
GBB4-Herzbeschwerden	Weiblich	32	4,625	4,556	2,781	2,649					
	Männlich	32	3,563	3,592	3,125	3,045					
	Gesamt	64	4,094	4,105	2,953	2,836	0,191	0,441	0,651	0,005	0,078
GBB5-Beschwerdedruck	Weiblich	31	22,484	15,347	18,742	14,429					
	Männlich	31	20,290	10,799	18,935	10,276					
	Gesamt	62	21,387	13,206	18,839	12,423	0,059	0,068	0,745	0,035	0,316
MMPI-Hypochondriasis	Weiblich	31	7,258	5,170	5,871	4,801					
	Männlich	32	7,469	4,142	6,719	3,879					
	Gesamt	63	7,365	4,640	6,302	4,343	0,223	0,241	0,617	0,018	0,469
MMPI-Depression	Weiblich	31	10,194	5,753	9,161	6,246					
	Männlich	33	11,485	4,086	10,061	3,905					
	Gesamt	64	10,859	4,966	9,625	5,150	0,061	0,011	0,340	0,031	0,726
MMPI-Hysteria	Weiblich	29	11,310	6,788	9,931	5,806					
	Männlich	32	13,063	4,996	11,375	5,021					
	Gesamt	61	12,230	5,931	10,689	5,412	0,099	0,431	0,232	0,013	0,798
MMPI-Psychasthenia	Weiblich	31	11,419	5,427	8,839	5,580					
	Männlich	32	12,937	5,588	11,625	4,818					
	Gesamt	63	12,190	5,518	10,254	5,352	0,872	0,418	0,090	0,000	0,226



**Tabelle 62b: Varianzanalyse (Manova): Therapeutengeschlecht, Zusammenhang mit den MMPI(K)-Skalen L, V, C, PD, M/F, PA, SC, HA, SI und VEV im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Therapeutengeschlecht		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Weiblich	29	3,517	1,617	4,000	1,711					
	Männlich	30	3,533	1,479	3,333	1,373					
	Gesamt	59	3,525	1,535	3,661	1,571	0,634	0,005	0,363	0,465	0,081
MMPI-Validity	Weiblich	25	3,240	2,505	2,920	2,120					
	Männlich	29	4,517	3,915	4,276	3,127					
	Gesamt	54	3,926	3,369	3,648	2,769	0,026	0,241	0,081	0,461	0,918
MMPI-Correction	Weiblich	29	9,448	3,651	10,897	3,904					
	Männlich	32	10,031	3,797	9,906	3,701					
	Gesamt	61	9,754	3,709	10,377	3,800	0,831	0,052	0,822	0,063	0,028
MMPI-Psychopathic Deviate	Weiblich	31	9,355	4,371	8,000	4,612					
	Männlich	30	10,733	4,185	10,967	3,755					
	Gesamt	61	10,033	4,301	9,459	4,437	0,813	0,773	0,036	0,166	0,052
MMPI-Masculinity/Femininity	Weiblich	31	14,839	3,184	14,355	3,050					
	Männlich	32	14,500	3,341	14,500	3,943					
	Gesamt	63	14,667	3,243	14,429	3,504	0,791	0,266	0,907	0,281	0,281
MMPI-Paranoia	Weiblich	29	6,000	3,059	4,690	2,661					
	Männlich	32	7,125	3,957	6,750	3,445					
	Gesamt	61	6,590	3,575	5,770	3,242	0,164	0,158	0,049	0,011	0,150
MMPI-Schizophrenia	Weiblich	22	7,636	5,412	6,000	4,024					
	Männlich	31	7,419	4,566	6,613	3,547					
	Gesamt	53	7,509	4,886	6,358	3,727	0,392	0,163	0,858	0,024	0,433
MMPI-Hypomania	Weiblich	28	10,286	4,090	8,821	3,591					
	Männlich	31	9,774	3,232	10,290	3,635					
	Gesamt	59	10,017	3,641	9,593	3,658	0,211	0,524	0,595	0,137	0,003
MMPI-Social Introversion	Weiblich	30	13,067	4,690	11,600	5,481					
	Männlich	30	13,067	4,234	13,233	5,184					
	Gesamt	60	13,067	4,430	12,417	5,353	0,585	0,948	0,483	0,221	0,126
VEV	Weiblich	32	196,563	38,870	199,344	52,302					
	Männlich	30	197,067	30,118	209,176	31,003					
	Gesamt	62	196,806	34,629	204,097	43,262	0,168	0,766	0,514	0,232	0,453

**Tabelle 63a: Varianzanalyse (Manova): Therapeutengeschlecht, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Therapeutengeschlecht		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
GBB1-Erschöpfungsneigung	Weiblich	36	10,611	6,429	6,278	5,527					
	Männlich	37	11,514	5,491	7,135	4,590					
	Gesamt	73	11,068	5,947	6,712	5,057	0,348	0,269	0,436	0,000	0,972
GBB2-Magenbeschwerden	Weiblich	34	4,500	3,527	2,794	3,488					
	Männlich	35	6,400	5,001	3,486	3,100					
	Gesamt	69	5,464	4,411	3,145	3,291	0,045	0,495	0,106	0,000	0,216
GBB3-Gliederschmerzen	Weiblich	36	9,750	5,541	6,111	4,640					
	Männlich	35	7,971	5,142	5,571	4,533					
	Gesamt	71	8,873	5,385	5,845	4,563	0,661	0,891	0,258	0,000	0,311
GBB4-Herzbeschwerden	Weiblich	36	5,639	3,226	2,972	2,903					
	Männlich	35	6,543	4,617	3,314	3,864					
	Gesamt	71	6,085	3,970	3,141	3,390	0,037	0,095	0,376	0,000	0,599
GBB5-Beschwerdedruck	Weiblich	34	29,912	14,764	18,324	14,068					
	Männlich	31	31,419	16,029	19,258	12,837					
	Gesamt	65	30,631	15,277	18,769	13,399	0,645	0,607	0,694	0,000	0,877
MMPI-Hypochondriasis	Weiblich	33	8,909	4,585	5,455	4,790					
	Männlich	36	10,222	4,882	6,806	4,426					
	Gesamt	69	9,594	4,754	6,159	4,620	0,716	0,648	0,184	0,000	0,972
MMPI-Depression	Weiblich	35	13,714	5,496	8,257	6,409					
	Männlich	37	14,595	4,226	9,946	4,403					
	Gesamt	72	14,167	4,870	9,125	5,498	0,125	0,029	0,210	0,000	0,555
MMPI-Hysteria	Weiblich	31	14,484	5,458	9,419	5,915					
	Männlich	36	16,222	5,499	11,083	5,669					
	Gesamt	67	15,418	5,508	10,313	5,800	0,966	0,809	0,133	0,000	0,963
MMPI-Psychasthenia	Weiblich	34	14,176	5,807	8,206	5,714					
	Männlich	34	16,382	4,930	11,029	5,096					
	Gesamt	68	15,279	5,460	9,618	5,558	0,351	0,514	0,025	0,000	0,669

**Tabelle 63b: Varianzanalyse (Manova): Therapeutengeschlecht, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Therapeutengeschlecht		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Weiblich	32	3,594	2,256	4,031	1,787					
	Männlich	34	3,441	1,779	3,294	1,548					
	Gesamt	66	3,515	2,010	3,652	1,696	0,184	0,421	0,256	0,548	0,229
MMPI-Validity	Weiblich	29	5,138	3,989	2,897	1,988					
	Männlich	32	6,094	4,431	4,094	3,094					
	Gesamt	61	5,539	4,219	3,525	2,675	0,568	0,018	0,155	0,000	0,813
MMPI-Correction	Weiblich	32	8,906	3,930	11,031	3,806					
	Männlich	35	8,457	3,023	9,686	3,724					
	Gesamt	67	8,672	3,466	10,328	3,795	0,137	0,902	0,229	0,001	0,366
MMPI-Psychopathic Deviate	Weiblich	33	11,667	4,781	8,091	4,844					
	Männlich	32	13,500	3,810	11,000	3,637					
	Gesamt	65	12,569	4,394	9,523	4,504	0,205	0,110	0,008	0,000	0,399
MMPI-Masculinity/Femininity	Weiblich	32	14,812	3,402	14,625	3,250					
	Männlich	33	15,485	3,510	14,636	4,144					
	Gesamt	65	15,154	3,447	14,631	3,702	0,861	0,175	0,688	0,072	0,247
MMPI-Paranoia	Weiblich	33	7,879	4,204	5,000	2,537					
	Männlich	31	8,581	4,072	7,000	3,531					
	Gesamt	64	8,219	4,123	5,969	3,197	0,861	0,070	0,084	0,000	0,189
MMPI-Schizophrenia	Weiblich	26	9,538	5,171	5,423	3,962					
	Männlich	31	10,032	5,269	6,323	3,727					
	Gesamt	57	9,807	5,184	5,912	3,828	0,921	0,749	0,502	0,000	0,756
MMPI-Hypomania	Weiblich	31	10,548	3,414	9,000	3,540					
	Männlich	31	10,194	2,926	9,903	3,944					
	Gesamt	62	10,371	3,158	9,452	3,745	0,403	0,558	0,730	0,023	0,116
MMPI-Social Introversion	Weiblich	33	13,576	4,549	10,939	5,701					
	Männlich	34	14,735	3,934	13,206	5,192					
	Gesamt	67	14,164	4,255	12,090	5,526	0,408	0,594	0,091	0,002	0,402

**Tabelle 64a: Varianzanalyse (Manova): Therapeutenberuf, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Therapeutenberuf		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
GBB1-Erschöpfungsneigung	Arzt	14	12,786	5,265	9,786	7,149	0,606	0,010	0,070	0,000	0,810
	Psychologe	90	10,456	5,661	7,089	4,960					
	Gesamt	104	10,769	5,642	7,452	5,344					
GBB2-Magenbeschwerden	Arzt	14	4,643	3,954	2,786	2,694	0,397	0,058	0,332	0,000	0,999
	Psychologe	90	5,644	4,455	3,789	3,524					
	Gesamt	104	5,510	4,387	3,654	3,430					
GBB3-Gliederschmerzen	Arzt	15	7,867	5,475	6,133	4,926	0,686	0,157	0,887	0,000	0,463
	Psychologe	86	8,465	5,169	5,872	4,026					
	Gesamt	101	8,376	5,192	5,911	4,145					
GBB4-Herzbeschwerden	Arzt	14	5,929	3,050	4,357	4,378	0,166	0,238	0,733	0,001	0,652
	Psychologe	87	5,839	3,716	3,805	3,701					
	Gesamt	101	5,851	3,618	3,881	3,782					
GBB5-Beschwerdedruck	Arzt	12	30,083	14,706	23,833	17,750	0,731	0,005	0,576	0,000	0,346
	Psychologe	82	29,866	13,976	19,976	11,590					
	Gesamt	94	29,894	13,990	20,468	12,487					
MMPI-Hypochondriasis	Arzt	15	10,067	4,758	7,867	5,423	0,666	0,156	0,643	0,000	0,956
	Psychologe	87	9,506	4,474	7,368	4,431					
	Gesamt	102	9,588	4,497	7,441	4,563					
MMPI-Depression	Arzt	14	16,000	3,843	13,286	5,269	0,358	0,866	0,056	0,000	0,602
	Psychologe	89	13,978	4,374	10,607	5,145					
	Gesamt	103	14,252	4,345	10,971	5,218					
MMPI-Hysteria	Arzt	13	16,385	5,440	13,846	7,358	0,596	0,084	0,367	0,000	0,620
	Psychologe	85	15,447	5,037	12,141	5,720					
	Gesamt	98	15,571	5,073	12,367	5,947					
MMPI-Psychasthenia	Arzt	14	16,929	4,531	14,571	5,125	0,413	0,470	0,061	0,000	0,630
	Psychologe	88	14,670	5,092	11,648	5,679					
	Gesamt	102	14,980	5,058	12,049	5,673					

**Tabelle 64b: Varianzanalyse (Manova): Therapeutenberuf, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Therapeutenberuf		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Arzt	12	3,250	1,603	3,000	1,477	0,232	0,172	0,220	0,965	0,294
	Psychologe	81	3,617	1,914	3,889	1,810					
	Gesamt	93	3,570	1,873	3,774	1,788					
MMPI-Validity	Arzt	11	6,818	3,188	7,182	5,016	0,104	0,000	0,020	0,530	0,198
	Psychologe	77	5,078	4,097	4,026	2,865					
	Gesamt	88	5,295	4,021	4,420	3,342					
MMPI-Correction	Arzt	14	6,929	3,125	8,357	3,478	0,677	0,575	0,080	0,008	0,666
	Psychologe	87	8,724	3,316	9,759	3,766					
	Gesamt	101	8,475	3,333	9,564	3,743					
MMPI-Psychopathic Deviate	Arzt	13	15,308	2,626	13,923	3,968	0,000	0,381	0,005	0,003	0,522
	Psychologe	83	12,289	4,512	10,169	4,512					
	Gesamt	96	12,698	4,418	10,677	4,608					
MMPI-Masculinity/Femininity	Arzt	14	15,429	4,164	14,714	3,496	0,019	0,382	0,859	0,052	0,609
	Psychologe	84	15,128	2,966	14,709	3,083					
	Gesamt	100	15,170	3,137	14,710	3,125					
MMPI-Paranoia	Arzt	11	9,636	4,202	11,000	5,099	0,420	0,002	0,003	0,950	0,008
	Psychologe	83	7,747	3,728	6,446	3,186					
	Gesamt	94	7,968	3,811	6,979	3,730					
MMPI-Schizophrenia	Arzt	12	11,667	5,297	10,000	5,939	0,469	0,075	0,056	0,002	0,502
	Psychologe	83	9,566	4,763	7,000	4,567					
	Gesamt	95	9,832	4,855	7,379	4,829					
MMPI-Hypomania	Arzt	11	11,455	3,671	11,636	4,653	0,119	0,080	0,104	0,632	0,372
	Psychologe	83	10,217	2,909	9,614	3,585					
	Gesamt	94	10,362	3,012	9,851	3,753					
MMPI-Social Introversion	Arzt	15	15,800	3,932	14,600	4,548	0,401	0,778	0,123	0,033	0,918
	Psychologe	88	13,955	4,425	12,864	4,732					
	Gesamt	103	14,223	4,388	13,117	4,724					

**Tabelle 65a: Varianzanalyse (Manova): Therapeutenberuf, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Therapeutenberuf		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P			
GBB1-Erschöpfungsneigung	Arzt	11	9,545	7,285	8,455	5,716	0,039	0,565	0,276	0,265	0,619
	Psychologe	50	7,400	4,961	6,980	5,141					
	Gesamt	61	7,787	5,444	7,246	5,230					
GBB2-Magenbeschwerden	Arzt	11	2,364	2,580	1,636	1,567	0,165	0,000	0,086	0,184	0,825
	Psychologe	48	3,917	3,351	3,396	3,325					
	Gesamt	59	3,627	3,258	3,068	3,140					
GBB3-Gliederschmerzen	Arzt	11	6,091	5,612	4,636	3,906	0,094	0,536	0,510	0,259	0,246
	Psychologe	50	6,220	4,112	6,240	4,378					
	Gesamt	61	6,197	4,366	5,951	4,311					
GBB4-Herzbeschwerden	Arzt	10	4,100	4,383	2,100	2,558	0,800	0,437	0,576	0,007	0,423
	Psychologe	50	4,280	4,180	3,180	2,960					
	Gesamt	60	4,250	4,177	3,000	2,905					
GBB5-Beschwerdedruck	Arzt	10	21,700	18,583	16,400	10,906	0,033	0,345	0,661	0,030	0,339
	Psychologe	48	21,979	12,300	19,875	13,064					
	Gesamt	58	21,931	13,390	19,276	12,699					
MMPI-Hypochondriasis	Arzt	11	7,636	5,626	5,636	3,880	0,232	0,450	0,740	0,013	0,394
	Psychologe	48	7,604	4,495	6,604	4,471					
	Gesamt	59	7,610	4,672	6,424	4,352					
MMPI-Depression	Arzt	11	12,727	4,606	11,909	3,961	0,531	0,143	0,103	0,111	0,640
	Psychologe	49	10,531	5,181	9,041	5,220					
	Gesamt	60	10,933	5,115	9,567	5,107					
MMPI-Hysteria	Arzt	10	13,400	7,427	10,900	4,677	0,210	0,395	0,699	0,019	0,509
	Psychologe	47	12,149	5,842	10,723	5,506					
	Gesamt	57	12,368	6,093	10,754	5,329					
MMPI-Psychasthenia	Arzt	11	14,636	3,931	12,091	3,961	0,035	0,058	0,151	0,001	0,684
	Psychologe	48	11,875	5,859	9,875	5,670					
	Gesamt	59	12,390	5,626	10,288	5,433					

**Tabelle 65b: Varianzanalyse (Manova): Therapeutenberuf, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI und VEV im Zweijahreszeitraum von t3 bis t5 in der Untersuchungsgruppe**

	Therapeutenberuf	t3			t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Arzt	8	3,125	1,458	4,000	1,069	0,807	0,052	0,931	0,108	0,164
	Psychologe	47	3,574	1,529	3,638	1,566					
	Gesamt	55	3,509	1,514	3,691	1,502					
MMPI-Validity	Arzt	9	6,889	5,134	4,778	3,153	0,002	0,485	0,021	0,034	0,043
	Psychologe	41	3,488	2,656	3,439	2,730					
	Gesamt	50	4,100	3,436	3,680	2,824					
MMPI-Correction	Arzt	10	8,200	3,190	6,700	2,452	0,389	0,011	0,135	0,143	0,665
	Psychologe	47	9,830	3,761	10,745	4,013					
	Gesamt	57	9,544	3,694	10,386	3,849					
MMPI-Psychopathic Deviate	Arzt	10	13,300	3,683	12,400	4,088	0,449	0,844	0,005	0,194	0,737
	Psychologe	47	9,191	4,256	8,660	4,244					
	Gesamt	57	9,912	4,421	9,316	4,421					
MMPI-Masculinity/Femininity	Arzt	10	14,100	3,247	14,700	4,111	0,928	0,396	0,847	0,723	0,103
	Psychologe	49	14,816	3,193	14,429	3,506					
	Gesamt	59	14,695	3,185	14,475	3,578					
MMPI-Paranoia	Arzt	10	9,500	4,813	7,500	4,062	0,018	0,142	0,017	0,003	0,125
	Psychologe	47	6,191	3,041	5,532	3,063					
	Gesamt	57	6,772	3,596	5,877	3,306					
MMPI-Schizophrenia	Arzt	9	9,889	5,689	7,889	3,180	0,381	0,364	0,111	0,025	0,578
	Psychologe	41	7,146	4,746	5,927	3,837					
	Gesamt	50	7,640	4,981	6,280	3,774					
MMPI-Hypomania	Arzt	10	12,900	3,479	12,200	1,874	0,952	0,001	0,005	0,241	0,725
	Psychologe	45	9,444	3,520	9,067	3,744					
	Gesamt	55	10,073	3,731	9,636	3,674					
MMPI-Social Introversion	Arzt	10	14,600	3,777	14,900	4,175	0,275	0,150	0,138	0,632	0,375
	Psychologe	46	12,891	4,673	11,891	5,531					
	Gesamt	56	13,196	4,542	12,429	5,407					
VEV	Arzt	10	180,000	55,084	182,000	68,352	0,002	0,001	0,021	0,526	0,686
	Psychologe	48	210,104	29,750	210,083	35,008					
	Gesamt	58	197,466	35,687	205,241	43,160					

**Tabelle 66a: Varianzanalyse (Manova): Therapeutenberuf, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Therapeutenberuf		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
GBB1-Erschöpfungsneigung	Arzt	13	12,385	6,462	7,615	5,738	0,689	0,450	0,416	0,000	0,752
	Psychologe	56	10,875	6,030	6,661	5,035					
	Gesamt	69	11,159	6,094	6,841	5,144					
GBB2-Magenbeschwerden	Arzt	11	3,727	2,936	1,636	1,567	0,014	0,000	0,060	0,002	0,798
	Psychologe	54	5,963	4,584	3,519	3,517					
	Gesamt	65	5,585	4,412	3,200	3,336					
GBB3-Gliederschmerzen	Arzt	13	8,538	6,839	4,231	3,723	0,085	0,166	0,272	0,000	0,383
	Psychologe	54	9,296	5,046	6,407	4,752					
	Gesamt	67	9,149	5,389	5,985	4,627					
GBB4-Herzbeschwerden	Arzt	13	6,000	3,894	2,615	2,468	0,863	0,025	0,572	0,000	0,783
	Psychologe	54	6,333	4,014	3,333	3,676					
	Gesamt	67	6,269	3,964	3,194	3,470					
GBB5-Beschwerdedruck	Arzt	11	27,727	17,223	16,273	11,109	0,477	0,181	0,345	0,000	0,872
	Psychologe	50	32,100	15,105	19,820	14,231					
	Gesamt	61	31,311	15,448	19,180	13,706					
MMPI-Hypochondriasis	Arzt	13	9,385	4,753	5,308	3,637	0,922	0,116	0,535	0,000	0,584
	Psychologe	53	9,811	4,836	6,491	4,822					
	Gesamt	66	9,727	4,786	6,258	4,612					
MMPI-Depression	Arzt	13	14,923	4,699	10,385	5,205	0,677	0,725	0,348	0,000	0,668
	Psychologe	55	14,036	5,055	8,727	5,536					
	Gesamt	68	14,206	4,967	9,044	5,476					
MMPI-Hysteria	Arzt	13	15,308	5,574	10,000	4,564	0,975	0,106	0,836	0,000	0,948
	Psychologe	52	15,538	5,543	10,365	6,097					
	Gesamt	65	15,492	5,506	10,292	5,793					
MMPI-Psychasthenia	Arzt	12	16,083	4,078	10,167	5,306	0,052	0,701	0,607	0,000	0,837
	Psychologe	54	15,111	5,784	9,593	5,682					
	Gesamt	66	15,288	5,499	9,697	5,580					



**Tabelle 66b: Varianzanalyse (Manova): Therapeutenberuf, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Therapeutenberuf		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Arzt	12	3,000	1,859	3,833	1,642	0,534	0,945	0,638	0,198	0,157
	Psychologe	50	3,680	2,084	3,640	1,663					
	Gesamt	62	3,548	2,046	3,677	1,647					
MMPI-Validity	Arzt	9	7,000	2,915	4,667	3,202	0,027	0,275	0,192	0,003	0,952
	Psychologe	49	5,531	4,477	3,286	2,598					
	Gesamt	58	5,759	4,285	3,500	2,716					
MMPI-Correction	Arzt	11	7,182	3,093	9,364	2,942	0,517	0,093	0,163	0,005	0,716
	Psychologe	52	8,846	3,478	10,538	3,997					
	Gesamt	63	8,556	3,449	10,333	3,840					
MMPI-Psychopathic Deviate	Arzt	12	15,083	2,392	10,917	5,178	0,000	0,304	0,032	0,000	0,494
	Psychologe	50	11,980	4,684	8,940	4,283					
	Gesamt	62	12,581	4,492	9,323	4,493					
MMPI-Masculinity/Femininity	Arzt	10	15,000	3,494	14,700	4,111	0,976	0,604	0,932	0,252	0,680
	Psychologe	52	15,269	3,515	14,635	3,737					
	Gesamt	62	15,226	3,485	14,645	3,764					
MMPI-Paranoia	Arzt	10	9,700	4,322	7,000	4,243	0,769	0,076	0,211	0,001	0,782
	Psychologe	51	8,137	4,094	5,824	3,051					
	Gesamt	61	8,393	4,136	6,016	3,263					
MMPI-Schizophrenia	Arzt	9	9,222	4,893	6,556	3,909	0,662	0,972	0,973	0,000	0,435
	Psychologe	47	9,872	5,327	5,809	3,882					
	Gesamt	56	9,768	5,222	5,929	3,861					
MMPI-Hypomania	Arzt	11	11,364	3,042	10,364	4,032	0,749	0,663	0,288	0,085	0,877
	Psychologe	49	10,163	3,230	9,327	3,716					
	Gesamt	60	10,383	3,205	9,517	3,762					
MMPI-Social Introversion	Arzt	11	16,455	3,751	14,273	4,474	0,534	0,177	0,052	0,019	0,988
	Psychologe	52	13,769	4,199	11,615	5,719					
	Gesamt	63	14,238	4,222	12,079	5,583					

**Tabelle 67a: Varianzanalyse (Manova): Schulabschluß zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Schulabschluß		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
GBB1-Erschöpfungsneigung	Hauptschule	9	10,000	4,213	6,667	5,123	0,513	0,550	0,774	0,000	0,973
	Mittlere Reife	25	11,320	6,019	7,880	5,093					
	Abitur oder Studium	54	10,500	6,009	7,370	5,880					
	Gesamt	88	10,682	5,814	7,443	5,546					
GBB2-Magenbeschwerden	Hauptschule	9	2,778	1,986	2,889	2,315	0,007	0,434	0,329	0,003	0,153
	Mittlere Reife	24	5,750	3,467	3,917	3,387					
	Abitur oder Studium	55	5,182	4,418	3,400	3,424					
	Gesamt	88	5,091	4,042	3,489	3,301					
GBB3-Gliederschmerzen	Hauptschule	9	7,000	2,646	4,000	3,082	0,163	0,078	0,086	0,000	0,294
	Mittlere Reife	25	10,200	5,575	6,640	3,740					
	Abitur oder Studium	53	7,264	5,527	5,340	4,653					
	Gesamt	87	8,080	5,444	5,575	4,302					
GBB4-Herzbeschwerden	Hauptschule	9	6,556	5,003	3,778	3,866	0,007	0,883	0,386	0,000	0,649
	Mittlere Reife	24	6,208	3,413	4,000	3,563					
	Abitur oder Studium	53	5,000	2,968	3,302	4,036					
	Gesamt	86	5,500	3,363	3,547	3,861					
GBB5-Beschwerdedruck	Hauptschule	9	26,333	6,727	17,333	11,347	0,075	0,132	0,363	0,000	0,851
	Mittlere Reife	22	32,591	13,738	22,182	10,477					
	Abitur oder Studium	50	27,780	14,896	19,200	14,426					
	Gesamt	81	28,926	13,973	19,802	13,101					
MMPI-Hypochondriasis	Hauptschule	9	8,889	2,759	7,556	5,028	0,141	0,462	0,382	0,000	0,612
	Mittlere Reife	23	10,261	4,180	7,870	3,533					
	Abitur oder Studium	53	8,434	4,631	6,830	4,906					
	Gesamt	85	8,976	4,383	7,188	4,560					
MMPI-Depression	Hauptschule	8	11,625	4,307	10,000	4,276	0,965	0,063	0,509	0,000	0,276
	Mittlere Reife	26	14,885	4,150	10,846	4,183					
	Abitur oder Studium	53	13,679	4,535	10,981	5,826					
	Gesamt	87	13,851	4,445	10,851	5,213					
MMPI-Hysteria	Hauptschule	8	14,375	3,701	11,875	5,817	0,586	0,297	0,403	0,000	0,462
	Mittlere Reife	25	16,360	4,982	12,760	4,816					
	Abitur oder Studium	48	14,000	4,895	11,875	6,584					
	Gesamt	81	14,765	4,887	12,148	5,963					
MMPI-Psychasthenia	Hauptschule	9	14,333	6,708	10,333	5,500	0,055	0,311	0,670	0,000	0,230
	Mittlere Reife	23	15,783	4,317	12,130	4,414					
	Abitur oder Studium	53	14,151	5,238	12,189	6,099					
	Gesamt	85	14,612	5,164	11,976	5,598					

**Tabelle 67b: Varianzanalyse (Manova): Schulabschluß zum Zeitpunkt t1, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Schulabschluß	N	t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
			M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Hauptschule/weniger	9	4,111	0,928	5,444	1,333	0,005	0,361	0,089	0,031	0,097
	Mittlere Reife	23	3,739	2,320	3,652	1,898					
	Abitur und mehr	48	3,312	1,800	3,583	1,773					
	Gesamt	80	3,525	1,896	3,813	1,842					
MMPI-Validity	Hauptschule/weniger	5	4,600	3,362	5,000	3,674	0,200	0,615	0,553	0,294	0,537
	Mittlere Reife	22	6,000	4,451	4,545	3,814					
	Abitur und mehr	46	4,826	3,504	3,978	2,910					
	Gesamt	73	5,164	3,793	4,219	3,224					
MMPI-Correction	Hauptschule/weniger	8	9,625	1,923	12,000	4,408	0,128	0,367	0,118	0,002	0,337
	Mittlere Reife	23	7,696	3,197	8,913	3,884					
	Abitur und mehr	55	9,109	3,457	9,800	3,450					
	Gesamt	86	8,779	3,316	9,767	3,706					
MMPI-Psychopathic Deviate	Hauptschule/weniger	8	11,000	6,071	8,875	3,357	0,080	0,156	0,586	0,000	0,112
	Mittlere Reife	23	13,348	4,706	9,957	3,914					
	Abitur und mehr	49	12,061	4,140	10,633	4,751					
	Gesamt	80	12,325	4,514	10,263	4,391					
MMPI-Masculinity/Femininity	Hauptschule/weniger	9	13,000	4,500	14,111	3,951	0,023	0,228	0,513	0,999	0,062
	Mittlere Reife	24	15,042	3,210	14,500	3,176					
	Abitur und mehr	51	15,098	2,893	14,529	3,107					
	Gesamt	84	14,857	3,204	14,476	3,183					
MMPI-Paranoia	Hauptschule/weniger	6	7,667	3,933	6,833	2,639	0,760	0,340	0,170	0,023	0,469
	Mittlere Reife	22	9,000	4,209	7,227	3,477					
	Abitur und mehr	50	6,880	3,618	6,060	3,771					
	Gesamt	78	7,538	3,880	6,449	3,617					
MMPI-Schizophrenia	Hauptschule/weniger	7	8,714	3,988	5,571	2,370	0,865	0,000	0,641	0,000	0,235
	Mittlere Reife	22	9,864	4,673	6,409	3,112					
	Abitur und mehr	50	9,460	4,752	7,800	5,295					
	Gesamt	79	9,506	4,624	7,215	4,615					
MMPI-Hypomania	Hauptschule/weniger	8	9,375	3,335	6,625	3,543	0,877	0,975	0,161	0,003	0,004
	Mittlere Reife	23	10,913	3,260	9,826	3,857					
	Abitur und mehr	47	9,957	2,836	10,319	3,636					
	Gesamt	78	10,179	3,018	9,795	3,808					
MMPI-Sozial Introversion	Hauptschule/weniger	9	12,778	4,631	11,778	4,295	0,667	0,249	0,138	0,014	0,173
	Mittlere Reife	24	16,042	3,605	13,917	3,955					
	Abitur und mehr	54	13,389	4,672	12,889	5,131					
	Gesamt	87	14,057	4,525	13,057	4,746					

**Tabelle 68a: Varianzanalyse (Manova): Schulabschluß zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Schulabschluß		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
GBB1-Erschöpfungsneigung	Hauptschule	5	6,800	4,438	5,400	2,608					
	Mittlere Reife	16	7,938	4,823	7,375	5,500					
	Abitur oder Studium	38	7,816	6,044	7,105	5,559					
	Gesamt	59	7,763	5,547	7,034	5,317	0,242	0,265	0,833	0,225	0,921
GBB2-Magenbeschwerden	Hauptschule	5	3,000	2,915	3,200	3,271					
	Mittlere Reife	14	3,357	3,433	3,071	3,872					
	Abitur oder Studium	38	3,789	3,354	3,026	3,027					
	Gesamt	57	3,614	3,294	3,053	3,209	0,949	0,430	0,963	0,568	0,689
GBB3-Gliederschmerzen	Hauptschule	5	4,600	3,362	4,600	2,966					
	Mittlere Reife	16	6,563	3,794	6,750	4,683					
	Abitur oder Studium	38	6,026	4,857	5,526	4,336					
	Gesamt	59	6,051	4,451	5,780	4,323	0,164	0,331	0,565	0,882	0,830
GBB4-Herzbeschwerden	Hauptschule	5	3,600	4,506	2,600	3,435					
	Mittlere Reife	16	4,313	3,439	3,938	3,130					
	Abitur oder Studium	38	4,211	4,563	2,684	2,642					
	Gesamt	59	4,186	4,216	3,017	2,850	0,726	0,544	0,734	0,106	0,502
GBB5-Beschwerdedruck	Hauptschule	5	18,000	13,077	15,800	11,054					
	Mittlere Reife	14	22,857	10,458	21,643	13,171					
	Abitur oder Studium	38	21,842	14,802	18,342	12,887					
	Gesamt	57	21,754	13,562	18,930	12,715	0,374	0,926	0,695	0,193	0,741
MMPI-Hypochondriasis	Hauptschule	5	8,600	5,128	6,400	4,099					
	Mittlere Reife	15	7,267	3,127	7,333	3,994					
	Abitur oder Studium	37	7,135	5,271	5,757	4,687					
	Gesamt	57	7,298	4,728	6,228	4,448	0,397	0,663	0,747	0,074	0,322
MMPI-Depression	Hauptschule	5	10,400	3,647	8,400	3,647					
	Mittlere Reife	16	10,938	3,586	10,438	5,138					
	Abitur oder Studium	37	10,865	5,813	9,378	5,377					
	Gesamt	58	10,845	5,067	9,586	5,147	0,023	0,495	0,845	0,119	0,729
MMPI-Hysteria	Hauptschule	5	13,000	4,848	10,200	5,119					
	Mittlere Reife	16	11,875	3,649	11,188	4,636					
	Abitur oder Studium	35	12,000	6,979	10,257	5,863					
	Gesamt	56	12,054	5,962	10,518	5,404	0,019	0,514	0,959	0,053	0,647
MMPI-Psychasthenia	Hauptschule	5	11,800	5,495	10,200	4,764					
	Mittlere Reife	15	13,000	4,243	11,800	5,821					
	Abitur oder Studium	37	11,865	5,851	9,703	5,206					
	Gesamt	57	12,158	5,378	10,298	5,325	0,543	0,632	0,566	0,038	0,759

**Tabelle 68b: Varianzanalyse (Manova): Schulabschluß zum Zeitpunkt t1, Zusammenhang mit den MMPI(K)-Skalen L, V, C, PD, M/F, PA, SC, HA, SI und VEV im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Schulabschluß		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Hauptschule	5	4,600	0,894	4,000	1,581					
	Mittlere Reife	14	3,500	1,286	3,500	1,506					
	Abitur	34	3,412	1,708	3,735	1,729					
	Gesamt	53	3,547	1,564	3,698	1,636	0,047	0,838	0,526	0,747	0,429
MMPI-Validity	Hauptschule	3	3,333	3,215	5,000	4,583					
	Mittlere Reife	12	4,917	4,502	2,917	1,564					
	Abitur	34	3,441	2,776	3,412	2,176					
	Gesamt	49	3,796	3,285	3,388	2,216	0,087	0,000	0,775	0,832	0,032
MMPI-Correction	Hauptschule	5	11,200	5,357	11,600	4,159					
	Mittlere Reife	15	8,933	3,882	9,600	3,699					
	Abitur	36	9,861	3,514	10,444	3,684					
	Gesamt	56	9,732	3,763	10,321	3,698	0,087	0,791	0,463	0,303	0,984
MMPI-Psychopathic Deviate	Hauptschule	5	9,000	2,550	8,200	3,701					
	Mittlere Reife	16	10,000	3,141	9,625	4,225					
	Abitur	36	10,111	4,803	9,444	4,359					
	Gesamt	57	9,982	4,198	9,386	4,216	0,012	0,925	0,812	0,309	0,949
MMPI-Masculinity/Femininity	Hauptschule	5	14,800	4,970	14,400	5,941					
	Mittlere Reife	16	15,125	2,849	15,125	2,849					
	Abitur	37	14,108	3,187	13,811	3,256					
	Gesamt	58	14,448	3,240	14,224	3,413	0,016	0,001	0,478	0,473	0,836
MMPI-Paranoia	Hauptschule	5	7,400	2,510	6,400	3,050					
	Mittlere Reife	15	7,267	3,918	6,067	2,685					
	Abitur	35	5,857	3,482	5,143	2,881					
	Gesamt	55	6,382	3,546	5,509	2,834	0,301	0,948	0,337	0,041	0,822
MMPI-Schizophrenia	Hauptschule	5	5,400	2,510	5,400	4,827					
	Mittlere Reife	13	7,077	3,303	7,000	3,786					
	Abitur	30	7,933	5,489	6,333	3,754					
	Gesamt	48	7,437	4,749	6,417	3,814	0,005	0,364	0,658	0,420	0,383
MMPI-Hypomania	Hauptschule	5	7,600	4,336	9,000	6,519					
	Mittlere Reife	16	10,375	4,145	9,688	3,321					
	Abitur	33	10,030	3,312	9,394	3,230					
	Gesamt	54	9,907	3,672	9,444	3,559	0,690	0,001	0,615	0,955	0,219
MMPI-Sozial Introversion	Hauptschule	5	11,800	5,675	12,800	5,675					
	Mittlere Reife	15	14,133	3,543	14,200	4,144					
	Abitur	35	12,971	4,541	11,714	5,717					
	Gesamt	55	13,182	4,368	12,491	5,350	0,147	0,710	0,399	0,934	0,385
VEV	Hauptschule	5	193,800	29,158	205,400	36,576					
	Mittlere Reife	16	197,187	31,111	188,250	58,112					
	Abitur	36	196,000	37,313	208,833	36,502					
	Gesamt	57	196,140	34,510	202,754	43,788	0,406	0,026	0,576	0,570	0,352

**Tabelle 69a: Varianzanalyse (Manova): Schulabschluß zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Schulabschluß		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
			N	M	s	M	s	COCH P			
GBB1-Erschöpfungsneigung	Hauptschule	8	7,125	4,612	4,750	3,655	0,394	0,523	0,224	0,000	0,560
	Mittlere Reife	18	11,778	6,413	6,833	5,404					
	Abitur oder Studium	40	11,025	5,807	6,975	5,456					
	Gesamt	66	10,758	5,936	6,667	5,240					
GBB2-Magenbeschwerden	Hauptschule	7	4,000	3,559	4,286	5,187	0,195	0,031	0,894	0,027	0,126
	Mittlere Reife	16	4,750	3,550	3,125	3,722					
	Abitur oder Studium	39	5,795	4,703	3,026	2,987					
	Gesamt	62	5,323	4,307	3,194	3,430					
GBB3-Gliederschmerzen	Hauptschule	8	8,375	4,173	6,625	6,632	0,575	0,036	0,372	0,000	0,364
	Mittlere Reife	17	10,706	5,698	6,176	4,489					
	Abitur oder Studium	39	8,026	5,460	5,333	4,245					
	Gesamt	64	8,781	5,435	5,719	4,596					
GBB4-Herzbeschwerden	Hauptschule	8	7,500	6,188	4,750	6,319	0,001	0,000	0,149	0,000	0,870
	Mittlere Reife	17	6,471	3,243	3,412	3,063					
	Abitur oder Studium	39	5,231	3,320	2,821	2,827					
	Gesamt	64	5,844	3,776	3,219	3,471					
GBB5-Beschwerdedruck	Hauptschule	7	24,143	12,389	21,429	21,540	0,968	0,014	0,804	0,000	0,216
	Mittlere Reife	14	33,143	14,883	18,929	12,970					
	Abitur oder Studium	37	29,378	14,902	18,108	12,970					
	Gesamt	58	29,655	14,619	18,707	13,950					
MMPI-Hypochondriasis	Hauptschule	8	8,250	5,007	7,125	6,312	0,839	0,106	0,691	0,000	0,304
	Mittlere Reife	17	10,353	4,197	6,353	4,329					
	Abitur oder Studium	37	8,865	4,803	5,703	4,696					
	Gesamt	62	9,194	4,655	6,065	4,770					
MMPI-Depression	Hauptschule	8	10,375	5,755	8,750	6,089	0,408	0,991	0,349	0,000	0,227
	Mittlere Reife	18	15,222	4,772	9,333	5,801					
	Abitur oder Studium	39	14,026	4,676	9,000	5,492					
	Gesamt	65	13,908	4,971	9,062	5,562					
MMPI-Hysteria	Hauptschule	8	12,625	6,346	10,500	7,964	0,440	0,093	0,685	0,000	0,362
	Mittlere Reife	18	16,278	4,909	10,167	5,316					
	Abitur oder Studium	34	14,676	5,487	10,118	5,840					
	Gesamt	60	14,883	5,465	10,183	5,896					
MMPI-Psychasthenia	Hauptschule	8	13,375	6,653	9,875	5,249	0,373	0,508	0,747	0,000	0,658
	Mittlere Reife	17	15,824	5,235	10,059	6,388					
	Abitur oder Studium	36	14,694	5,466	9,306	5,424					
	Gesamt	61	14,836	5,520	9,590	5,602					

**Tabelle 69b: Varianzanalyse (Manova): Schulabschluß zum Zeitpunkt t1, Zusammenhang mit den MMPI(K)-Skalen  
L, V; C, PD, M/F, PA, SC, HA, SI im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Schulabschluß		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
			N	M	s	M	s	COCH P			
MMPI-Lie	Hauptschule	8	4,125	1,553	4,000	2,000	0,081	0,540	0,724	0,931	0,315
	Mittlere Reife	17	3,824	2,506	3,412	1,502					
	Abitur	35	3,314	1,906	3,771	1,848					
	Gesamt	60	3,567	2,045	3,700	1,759					
MMPI-Validity	Hauptschule	5	3,000	2,646	4,600	3,578	0,294	0,002	0,395	0,042	0,003
	Mittlere Reife	16	7,125	4,193	2,875	1,408					
	Abitur	33	4,636	3,740	3,333	2,217					
	Gesamt	54	5,222	3,965	3,315	2,170					
MMPI-Correction	Hauptschule	8	9,625	2,200	10,375	4,173	0,469	0,789	0,450	0,030	0,730
	Mittlere Reife	16	7,750	3,606	9,813	3,746					
	Abitur	37	9,189	3,612	10,541	3,679					
	Gesamt	61	8,869	3,481	10,328	3,709					
MMPI-Psychopathic Deviate	Hauptschule	6	9,333	5,854	8,833	3,656	0,126	0,548	0,358	0,007	0,324
	Mittlere Reife	17	13,588	4,184	9,529	4,155					
	Abitur	36	12,222	4,250	9,583	4,638					
	Gesamt	59	12,322	4,482	9,492	4,352					
MMPI-Masculinity/Femininity	Hauptschule	7	13,000	4,243	13,857	5,336	0,427	0,012	0,239	0,894	0,082
	Mittlere Reife	17	15,647	4,015	15,647	3,121					
	Abitur	35	14,800	2,816	13,800	3,315					
	Gesamt	59	14,831	3,400	14,339	3,575					
MMPI-Paranoia	Hauptschule	7	7,286	3,147	6,857	3,485	0,352	0,201	0,040	0,002	0,156
	Mittlere Reife	16	10,063	4,419	6,500	2,394					
	Abitur	35	7,029	3,839	5,057	2,796					
	Gesamt	58	7,897	4,098	5,672	2,837					
MMPI-Schizophrenia	Hauptschule	8	6,625	3,962	5,250	4,892	0,814	0,355	0,410	0,000	0,372
	Mittlere Reife	14	10,214	5,041	5,929	3,931					
	Abitur	30	9,500	4,918	5,933	3,741					
	Gesamt	52	9,250	4,874	5,827	3,904					
MMPI-Hypomania	Hauptschule	8	9,000	2,204	8,500	5,928	0,054	0,005	0,248	0,080	0,565
	Mittlere Reife	17	11,471	3,777	9,941	3,344					
	Abitur	30	9,733	2,728	9,167	3,228					
	Gesamt	55	10,164	3,114	9,309	3,701					
MMPI-Sozial Introversion	Hauptschule	7	11,857	4,741	12,571	6,161	0,449	0,840	0,184	0,116	0,284
	Mittlere Reife	18	16,167	3,330	13,000	5,402					
	Abitur	36	13,333	4,395	11,611	5,668					
	Gesamt	61	14,000	4,344	12,131	5,587					

**Tabelle 70a: Varianzanalyse (Manova): Anfangs-Lie zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS,D,HY,PT im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Anfangs-Lie		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
<b>GBB1-Erschöpfungsneigung</b>	Geringes Anfangs-Lie	60	10,017	5,768	7,250	4,990	0,437	0,594	0,405	0,000	0,102
	Hohes Anfangs-Lie	47	11,638	5,181	7,128	5,371					
	Gesamt	107	10,729	5,552	7,196	5,137					
<b>GBB2-Magenbeschwerden</b>	Geringes Anfangs-Lie	60	5,467	4,659	3,633	3,390	0,416	0,858	0,733	0,000	0,969
	Hohes Anfangs-Lie	47	5,213	4,165	3,404	3,475					
	Gesamt	107	5,355	4,430	3,533	3,413					
<b>GBB3-Gliederschmerzen</b>	Geringes Anfangs-Lie	59	7,949	5,390	6,203	4,413	0,996	0,023	0,811	0,000	0,015
	Hohes Anfangs-Lie	45	8,822	5,386	4,933	3,194					
	Gesamt	104	8,327	5,380	5,654	3,965					
<b>GBB4-Herzbeschwerden</b>	Geringes Anfangs-Lie	56	4,964	3,578	3,321	3,428	0,750	0,867	0,071	0,000	0,170
	Hohes Anfangs-Lie	48	6,521	3,421	3,937	3,509					
	Gesamt	104	5,683	3,575	3,606	3,462					
<b>GBB5-Beschwerdedruck</b>	Geringes Anfangs-Lie	54	27,611	15,040	19,963	11,956	0,118	0,721	0,493	0,000	0,046
	Hohes Anfangs-Lie	43	31,791	11,979	18,907	11,351					
	Gesamt	97	29,464	13,857	19,495	11,643					
<b>MMPI-Hypochondriasis</b>	Geringes Anfangs-Lie	56	9,161	5,077	6,804	4,466	0,017	0,709	0,387	0,000	0,828
	Hohes Anfangs-Lie	48	9,750	3,617	7,563	4,237					
	Gesamt	104	9,433	4,452	7,154	4,357					
<b>MMPI-Depression</b>	Geringes Anfangs-Lie	58	14,241	4,289	11,052	5,313	0,805	0,382	0,518	0,000	0,714
	Hohes Anfangs-Lie	48	13,875	4,144	10,375	4,702					
	Gesamt	106	14,075	4,208	10,745	5,033					
<b>MMPI-Hysteria</b>	Geringes Anfangs-Lie	55	14,600	5,209	11,491	5,511	0,480	0,558	0,098	0,000	0,718
	Hohes Anfangs-Lie	43	16,395	4,701	12,907	5,999					
	Gesamt	98	15,388	5,047	12,112	5,743					
<b>MMPI-Psychasthenia</b>	Geringes Anfangs-Lie	57	15,123	4,762	12,175	5,742	0,474	0,638	0,380	0,000	0,920
	Hohes Anfangs-Lie	46	14,348	5,267	11,304	5,374					
	Gesamt	103	14,777	4,984	11,786	5,571					



**Tabelle 70b: Varianzanalyse (Manova): Anfangs-Lie zum Zeitpunkt t1, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI im Einjahreszeitraum von t1 bis t3 in der Untersuchungsgruppe**

	Anfangs-Lie		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Geringes Anfangs-Lie	50	2,140	0,969	2,980	1,392	0,042	0,058	0,000	0,117	0,000
	Hohes Anfangs-Lie	48	5,042	1,304	4,687	1,835					
	Gesamt	98	3,561	1,850	3,816	1,830					
MMPI-Validity	Geringes Anfangs-Lie	54	5,185	3,387	4,111	2,950	0,017	0,076	0,555	0,025	0,521
	Hohes Anfangs-Lie	35	5,371	4,875	4,771	3,866					
	Gesamt	89	5,258	4,013	4,371	3,335					
MMPI-Correction	Geringes Anfangs-Lie	60	7,950	3,524	8,850	3,786	0,243	0,428	0,005	0,000	0,248
	Hohes Anfangs-Lie	44	9,364	2,989	10,977	3,386					
	Gesamt	104	8,548	3,367	9,750	3,757					
MMPI-Psychopathic Deviate	Geringes Anfangs-Lie	55	13,382	4,340	10,945	4,648	0,973	0,733	0,145	0,000	0,243
	Hohes Anfangs-Lie	42	11,714	4,318	10,190	4,424					
	Gesamt	97	12,660	4,387	10,619	4,545					
MMPI-Masculinity/Femininity	Geringes Anfangs-Lie	57	15,281	3,299	14,860	3,259	0,414	0,739	0,562	0,054	0,916
	Hohes Anfangs-Lie	45	14,911	2,937	14,533	3,109					
	Gesamt	102	15,118	3,135	14,716	3,182					
MMPI-Paranoia	Geringes Anfangs-Lie	54	7,944	3,833	6,685	3,425	0,970	0,167	0,968	0,004	0,361
	Hohes Anfangs-Lie	42	7,619	3,812	6,952	4,196					
	Gesamt	96	7,802	3,808	6,802	3,763					
MMPI-Schizophrenia	Geringes Anfangs-Lie	55	9,764	4,635	7,491	4,985	0,459	0,315	0,683	0,000	0,400
	Hohes Anfangs-Lie	40	9,775	5,166	6,750	4,301					
	Gesamt	95	9,768	4,839	7,179	4,699					
MMPI-Hypomania	Geringes Anfangs-Lie	53	11,019	2,664	10,792	3,591	0,205	0,934	0,003	0,039	0,207
	Hohes Anfangs-Lie	42	9,524	3,210	8,595	3,548					
	Gesamt	95	10,358	2,996	9,821	3,719					
MMPI-Social Introversion	Geringes Anfangs-Lie	60	14,517	4,515	13,033	5,172	0,920	0,125	0,457	0,003	0,318
	Hohes Anfangs-Lie	46	13,522	4,579	12,783	4,174					
	Gesamt	106	14,085	4,548	12,925	4,744					

**Tabelle 71a: Varianzanalyse (Manova): Anfangs-Lie zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Anfangs-Lie		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
<b>GBB1-Erschöpfungsneigung</b>	Geringes Anfangs-Lie	34	7,206	5,290	7,029	5,675	0,726	0,266	0,853	0,375	0,598
	Hohes Anfangs-Lie	29	7,690	4,965	7,000	4,637					
	Gesamt	63	7,429	5,108	7,016	5,182					
<b>GBB2-Magenbeschwerden</b>	Geringes Anfangs-Lie	33	2,970	2,801	2,697	2,856	0,161	0,325	0,275	0,112	0,407
	Hohes Anfangs-Lie	28	4,071	3,630	3,214	3,425					
	Gesamt	61	3,475	3,228	2,934	3,114					
<b>GBB3-Gliederschmerzen</b>	Geringes Anfangs-Lie	34	6,118	4,842	5,647	4,773	0,012	0,068	0,713	0,699	0,546
	Hohes Anfangs-Lie	29	5,483	3,055	5,586	3,418					
	Gesamt	63	5,825	4,098	5,619	4,171					
<b>GBB4-Herzbeschwerden</b>	Geringes Anfangs-Lie	33	3,515	3,906	2,758	2,750	0,529	0,709	0,558	0,006	0,496
	Hohes Anfangs-Lie	29	4,207	3,478	2,966	2,946					
	Gesamt	62	3,839	3,698	2,855	2,822					
<b>GBB5-Beschwerdedruck</b>	Geringes Anfangs-Lie	32	19,625	13,493	18,000	13,617	0,181	0,250	0,626	0,058	0,605
	Hohes Anfangs-Lie	28	21,679	10,488	18,857	10,967					
	Gesamt	60	20,583	12,127	18,400	12,355					
<b>MMPI-Hypochondriasis</b>	Geringes Anfangs-Lie	33	6,333	4,735	5,697	4,348	0,191	0,664	0,227	0,022	0,358
	Hohes Anfangs-Lie	28	7,964	3,717	6,500	4,014					
	Gesamt	61	7,082	4,341	6,066	4,183					
<b>MMPI-Depression</b>	Geringes Anfangs-Lie	33	10,879	5,583	10,030	5,753	0,040	0,175	0,461	0,048	0,616
	Hohes Anfangs-Lie	29	10,310	3,809	8,897	4,475					
	Gesamt	62	10,613	4,806	9,500	5,184					
<b>MMPI-Hysteria</b>	Geringes Anfangs-Lie	31	10,935	5,938	10,387	5,451	0,499	0,781	0,443	0,015	0,115
	Hohes Anfangs-Lie	28	12,893	5,231	10,429	5,174					
	Gesamt	59	11,864	5,652	10,407	5,276					
<b>MMPI-Psychasthenia</b>	Geringes Anfangs-Lie	33	11,485	5,880	10,030	5,514	0,274	0,793	0,672	0,001	0,042
	Hohes Anfangs-Lie	28	12,464	4,803	10,143	5,254					
	Gesamt	61	11,934	5,391	10,082	5,352					

**Tabelle 71b: Varianzanalyse (Manova): Anfangs-Lie, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI und VEV im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Anfangs-Lie		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
<b>MMPI-Lie</b>	Geringes Anfangs-Lie	29	2,793	1,292	3,069	1,668	0,567	0,159	0,000	0,556	0,445
	Hohes Anfangs-Lie	28	4,321	1,442	4,286	1,272					
	Gesamt	57	3,544	1,559	3,667	1,596					
<b>MMPI-Validity</b>	Geringes Anfangs-Lie	31	3,645	2,811	3,161	2,083	0,054	0,011	0,185	0,585	0,506
	Hohes Anfangs-Lie	21	4,429	4,166	4,476	3,516					
	Gesamt	52	3,962	3,407	3,692	2,797					
<b>MMPI-Correction</b>	Geringes Anfangs-Lie	33	8,879	4,021	9,697	3,933	0,099	0,695	0,060	0,199	0,332
	Hohes Anfangs-Lie	26	10,962	2,946	11,077	3,654					
	Gesamt	59	9,797	3,708	10,305	3,843					
<b>MMPI-Psychopathic Deviate</b>	Geringes Anfangs-Lie	32	10,063	4,464	9,344	4,469	0,741	0,962	0,968	0,192	0,700
	Hohes Anfangs-Lie	28	9,857	4,196	9,464	4,509					
	Gesamt	60	9,967	4,306	9,400	4,450					
<b>MMPI-Masculinity/Femininity</b>	Geringes Anfangs-Lie	33	14,545	3,289	14,394	3,418	0,987	0,680	0,890	0,214	0,562
	Hohes Anfangs-Lie	29	14,793	3,299	14,379	3,688					
	Gesamt	62	14,661	3,269	14,387	3,517					
<b>MMPI-Paranoia</b>	Geringes Anfangs-Lie	32	6,313	3,126	5,313	2,442	0,132	0,008	0,358	0,032	0,402
	Hohes Anfangs-Lie	27	6,815	4,151	6,370	4,059					
	Gesamt	59	6,542	3,607	5,797	3,295					
<b>MMPI-Schizophrenia</b>	Geringes Anfangs-Lie	30	7,033	4,832	6,167	3,887	0,816	0,831	0,689	0,055	0,788
	Hohes Anfangs-Lie	21	7,619	4,610	6,476	3,723					
	Gesamt	51	7,275	4,704	6,294	3,786					
<b>MMPI-Hypomania</b>	Geringes Anfangs-Lie	30	10,767	3,674	9,833	3,260	0,808	0,186	0,212	0,252	0,116
	Hohes Anfangs-Lie	27	9,074	3,507	9,222	4,200					
	Gesamt	57	9,965	3,664	9,544	3,713					
<b>MMPI-Social Introversion</b>	Geringes Anfangs-Lie	31	12,935	5,428	12,484	6,021	0,005	0,190	0,906	0,404	0,995
	Hohes Anfangs-Lie	27	13,074	3,149	12,630	4,683					
	Gesamt	58	13,000	4,476	12,552	5,394					
<b>VEV</b>	Geringes Anfangs-Lie	32	200,031	40,095	200,562	52,633	0,067	0,003	0,973	0,336	0,380
	Hohes Anfangs-Lie	28	194,821	28,360	206,321	29,948					
	Gesamt	60	197,600	34,923	203,250	43,294					

**Tabelle 72a: Varianzanalyse (Manova): Anfangs-Lie zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Anfangs-Lie		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
<b>GBB1-Erschöpfungsneigung</b>	Geringes Anfangs-Lie	38	10,447	6,008	6,605	5,524	0,716	0,295	0,628	0,000	0,488
	Hohes Anfangs-Lie	33	11,455	5,646	6,697	4,620					
	Gesamt	71	10,915	5,823	6,648	5,088					
<b>GBB2-Magenbeschwerden</b>	Geringes Anfangs-Lie	37	5,162	4,425	2,757	2,763	0,854	0,051	0,442	0,000	0,892
	Hohes Anfangs-Lie	30	5,733	4,571	3,467	3,902					
	Gesamt	67	5,418	4,466	3,075	3,313					
<b>GBB3-Gliederschmerzen</b>	Geringes Anfangs-Lie	37	8,189	5,734	5,730	4,647	0,413	0,757	0,541	0,000	0,235
	Hohes Anfangs-Lie	32	9,562	4,977	5,625	4,405					
	Gesamt	69	8,826	5,401	5,681	4,503					
<b>GBB4-Herzbeschwerden</b>	Geringes Anfangs-Lie	36	5,417	4,312	2,694	2,806	0,196	0,049	0,175	0,000	0,714
	Hohes Anfangs-Lie	33	6,576	3,446	3,455	3,954					
	Gesamt	69	5,971	3,937	3,058	3,399					
<b>GBB5-Beschwerdedruck</b>	Geringes Anfangs-Lie	34	28,794	16,753	17,794	13,223	0,149	0,832	0,491	0,000	0,623
	Hohes Anfangs-Lie	29	31,862	12,883	19,000	13,740					
	Gesamt	63	30,206	15,057	18,349	13,367					
<b>MMPI-Hypochondriasis</b>	Geringes Anfangs-Lie	35	8,571	5,359	5,343	4,359	0,057	0,759	0,117	0,000	0,572
	Hohes Anfangs-Lie	32	10,438	3,826	6,594	4,599					
	Gesamt	67	9,463	4,749	5,940	4,485					
<b>MMPI-Depression</b>	Geringes Anfangs-Lie	37	14,297	4,766	9,351	5,851	0,916	0,492	0,487	0,000	0,969
	Hohes Anfangs-Lie	33	13,606	4,854	8,606	5,196					
	Gesamt	70	13,971	4,785	9,000	5,525					
<b>MMPI-Hysteria</b>	Geringes Anfangs-Lie	35	14,200	5,645	9,457	5,570	0,556	0,802	0,120	0,000	0,680
	Hohes Anfangs-Lie	60	16,400	5,083	10,733	5,825					
	Gesamt	65	15,215	5,464	10,046	5,680					
<b>MMPI-Psychasthenia</b>	Geringes Anfangs-Lie	35	16,057	5,081	9,429	5,580	0,592	0,987	0,357	0,000	0,148
	Hohes Anfangs-Lie	31	13,935	5,591	9,452	5,597					
	Gesamt	66	15,061	5,392	9,439	5,545					

**Tabelle 72b: Varianzanalyse (Manova): Anfangs-Lie, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Anfangs-Lie		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Geringes Anfangs-Lie	34	1,914	1,127	2,882	1,629	0,623	0,328	0,000	0,615	0,000
	Hohes Anfangs-Lie	32	5,188	1,230	4,469	1,367					
	Gesamt	66	3,515	2,010	3,652	1,696					
MMPI-Validity	Geringes Anfangs-Lie	35	5,971	3,808	3,029	2,007	0,167	0,007	0,695	0,000	0,060
	Hohes Anfangs-Lie	24	5,292	4,939	4,333	3,371					
	Gesamt	59	5,695	4,276	3,559	2,699					
MMPI-Correction	Geringes Anfangs-Lie	36	7,417	3,516	9,722	3,859	0,284	0,884	0,008	0,003	0,129
	Hohes Anfangs-Lie	29	10,172	2,904	10,931	3,760					
	Gesamt	65	8,646	3,515	10,262	3,834					
MMPI-Psychopathic Deviate	Geringes Anfangs-Lie	35	13,371	4,045	9,229	4,346	0,378	0,601	0,513	0,000	0,075
	Hohes Anfangs-Lie	29	11,621	4,747	9,759	4,778					
	Gesamt	64	12,578	4,429	9,469	4,518					
MMPI-Masculinity/Femininity	Geringes Anfangs-Lie	34	15,412	3,543	14,824	3,605	0,852	0,676	0,536	0,065	0,879
	Hohes Anfangs-Lie	30	14,833	3,425	14,333	3,889					
	Gesamt	64	15,141	3,473	14,594	3,719					
MMPI-Paranoia	Geringes Anfangs-Lie	35	8,400	4,299	5,343	2,363	0,728	0,005	0,543	0,000	0,045
	Hohes Anfangs-Lie	27	7,889	4,032	6,852	4,007					
	Gesamt	62	8,177	4,159	6,000	3,244					
MMPI-Schizophrenia	Geringes Anfangs-Lie	31	9,968	5,174	5,774	3,922	0,818	0,977	0,819	0,000	0,567
	Hohes Anfangs-Lie	24	9,333	5,411	5,917	3,900					
	Gesamt	55	9,691	5,238	5,836	3,877					
MMPI-Hypomania	Geringes Anfangs-Lie	31	11,742	3,055	9,871	3,212	0,542	0,111	0,018	0,019	0,024
	Hohes Anfangs-Lie	29	8,931	2,725	8,897	4,337					
	Gesamt	60	10,383	3,205	9,400	3,797					
MMPI-Social Introversion	Geringes Anfangs-Lie	35	14,743	4,507	12,029	6,138	0,548	0,221	0,691	0,006	0,244
	Hohes Anfangs-Lie	30	13,533	4,049	12,400	4,931					
	Gesamt	65	14,185	4,312	12,200	5,574					

**Tabelle 73a : Varianzanalyse (Manova): Anfangsbeschwerdedruck zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Anfangsbeschwerdedruck	N	t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
			M	s	M	s	COCH P	COCH P	F	F	F
<b>GBB1-Erschöpfungsneigung</b>	Geringer Beschwerdedruck	52	6,962	3,865	5,519	3,622					
	Hoher Beschwerdedruck	50	14,620	4,472	9,080	6,372					
	Gesamt	102	10,716	5,661	7,265	5,434	0,306	0,000	0,000	0,000	0,000
<b>GBB2-Magenbeschwerden</b>	Geringer Beschwerdedruck	52	3,135	3,343	2,596	3,011					
	Hoher Beschwerdedruck	50	7,480	4,400	4,340	3,384					
	Gesamt	102	5,265	4,450	3,451	3,302	0,055	0,412	0,000	0,000	0,000
<b>GBB3-Gliederschmerzen</b>	Geringer Beschwerdedruck	52	4,808	3,453	4,308	3,178					
	Hoher Beschwerdedruck	49	12,184	4,653	7,592	4,458					
	Gesamt	101	8,386	5,495	5,901	4,173	0,037	0,018	0,000	0,000	0,000
<b>GBB4-Herzbeschwerden</b>	Geringer Beschwerdedruck	52	3,885	2,431	2,769	3,065					
	Hoher Beschwerdedruck	48	7,354	3,467	4,521	3,881					
	Gesamt	100	5,550	3,433	3,610	3,573	0,014	0,102	0,000	0,000	0,013
<b>GBB5-Beschwerdedruck</b>	Geringer Beschwerdedruck	52	18,788	6,921	15,192	7,802					
	Hoher Beschwerdedruck	47	41,894	9,015	25,362	14,328					
	Gesamt	99	29,758	14,055	20,020	12,413	0,067	0,000	0,000	0,000	0,000
<b>MMPI-Hypochondriasis</b>	Geringer Beschwerdedruck	51	6,725	3,317	5,392	3,505					
	Hoher Beschwerdedruck	48	12,229	3,657	8,917	4,712					
	Gesamt	99	9,394	4,435	7,101	4,478	0,498	0,041	0,000	0,000	0,011
<b>MMPI-Depression</b>	Geringer Beschwerdedruck	51	12,627	3,847	9,647	4,237					
	Hoher Beschwerdedruck	48	15,708	4,136	12,188	5,855					
	Gesamt	99	14,121	4,260	10,879	5,218	0,614	0,026	0,001	0,000	0,533
<b>MMPI-Hysteria</b>	Geringer Beschwerdedruck	49	12,878	4,265	10,163	4,625					
	Hoher Beschwerdedruck	43	18,279	4,511	14,721	6,580					
	Gesamt	92	15,402	5,131	12,293	6,041	0,709	0,020	0,000	0,000	0,446
<b>MMPI-Psychasthenia</b>	Geringer Beschwerdedruck	50	13,520	4,752	10,820	4,650					
	Hoher Beschwerdedruck	49	16,224	4,853	12,959	6,422					
	Gesamt	99	14,859	4,967	11,879	5,672	0,883	0,026	0,010	0,000	0,567

**Tabelle 73b: Varianzanalyse (Manova): Anfangsbeschwerdedruck zum Zeitpunkt t1, Zusammenhang mit den MMPI(K)-Skalen L, V, C, PD, M/F, PA, SC, HA, SI im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Anfangsbeschwerdedruck	t1			t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Geringer Beschwerdedruck	46	3,261	1,757	3,826	1,730					
	Höherer Beschwerdedruck	43	3,930	2,017	3,953	2,058					
	Gesamt	89	3,584	1,906	3,888	1,886	0,364	0,252	0,271	0,100	0,129
MMPI-Validity	Geringer Beschwerdedruck	44	4,114	2,599	3,386	2,535					
	Höherer Beschwerdedruck	39	6,231	4,749	5,103	3,169					
	Gesamt	83	5,108	3,889	4,193	2,961	0,000	0,158	0,004	0,012	0,580
MMPI-Correction	Geringer Beschwerdedruck	51	9,627	3,066	10,588	3,407					
	Höherer Beschwerdedruck	48	7,542	3,402	8,854	4,053					
	Gesamt	99	8,616	3,383	9,747	3,815	0,469	0,228	0,003	0,001	0,582
MMPI-Psychopathic Deviate	Geringer Beschwerdedruck	49	11,653	4,489	9,837	4,007					
	Höherer Beschwerdedruck	44	13,727	4,234	11,545	5,092					
	Gesamt	93	12,634	4,469	10,645	4,608	0,693	0,108	0,026	0,000	0,647
MMPI-Masculinity/Femininity	Geringer Beschwerdedruck	49	14,980	2,983	14,857	3,075					
	Höherer Beschwerdedruck	46	15,065	3,376	14,543	3,264					
	Gesamt	95	15,021	3,162	14,705	3,155	0,399	0,685	0,854	0,118	0,331
MMPI-Paranoia	Geringer Beschwerdedruck	47	6,553	2,999	5,511	2,970					
	Höherer Beschwerdedruck	43	8,907	4,185	7,791	3,732					
	Gesamt	90	7,678	3,783	6,600	3,528	0,029	0,133	0,001	0,002	0,912
MMPI-Schizophrenia	Geringer Beschwerdedruck	47	8,021	3,644	5,489	3,196					
	Höherer Beschwerdedruck	44	11,591	4,957	9,091	5,548					
	Gesamt	91	9,747	4,663	7,231	4,817	0,042	0,000	0,000	0,000	0,972
MMPI-Hypomania	Geringer Beschwerdedruck	45	9,778	3,125	9,489	3,935					
	Höherer Beschwerdedruck	44	10,955	2,736	10,250	3,590					
	Gesamt	89	10,360	2,982	9,865	3,766	0,381	0,546	0,141	0,099	0,488
MMPI-Social Introversion	Geringer Beschwerdedruck	51	13,627	4,395	12,588	4,691					
	Höherer Beschwerdedruck	49	13,959	4,444	13,633	4,720					
	Gesamt	100	13,790	4,400	13,100	4,711	0,939	0,966	0,417	0,052	0,308

**Tabelle 74a: Varianzanalyse (Manova): Anfangsbeschwerdedruck zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Anfangsbeschwerdedruck	N	t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
			M	s	M	s	COCH P	COCH P	F	F	F
<b>GBB1-Erschöpfungsneigung</b>	Geringer Beschwerdedruck	32	5,219	3,024	4,781	4,187					
	Hohe Beschwerdedruck	28	10,357	6,453	9,679	5,143					
	Gesamt	60	7,617	5,527	7,067	5,233	0,000	0,274	0,000	0,278	0,814
<b>GBB2-Magenbeschwerden</b>	Geringer Beschwerdedruck	31	2,387	2,654	2,097	2,844					
	Hohe Beschwerdedruck	28	4,607	3,213	3,964	3,237					
	Gesamt	59	3,441	3,114	2,983	3,154	0,309	0,490	0,005	0,188	0,616
<b>GBB3-Gliederschmerzen</b>	Geringer Beschwerdedruck	32	4,094	2,966	4,031	3,754					
	Hohe Beschwerdedruck	28	8,214	4,693	7,500	4,123					
	Gesamt	60	6,017	4,359	5,650	4,270	0,016	0,617	0,000	0,436	0,513
<b>GBB4-Herzbeschwerden</b>	Geringer Beschwerdedruck	32	2,594	2,872	2,250	2,383					
	Hohe Beschwerdedruck	27	5,259	4,338	3,037	2,766					
	Gesamt	59	3,814	3,826	2,610	2,573	0,030	0,426	0,019	0,002	0,022
<b>GBB5-Beschwerdedruck</b>	Geringer Beschwerdedruck	31	14,323	7,007	13,097	10,830					
	Hohe Beschwerdedruck	27	28,481	14,898	24,259	11,651					
	Gesamt	58	20,914	13,335	18,293	12,458	0,000	0,702	0,000	0,035	0,240
<b>MMPI-Hypochondriasis</b>	Geringer Beschwerdedruck	32	4,906	3,325	4,188	3,847					
	Hohe Beschwerdedruck	27	9,741	4,638	8,630	3,855					
	Gesamt	59	7,119	4,632	6,220	4,422	0,078	0,992	0,000	0,047	0,665
<b>MMPI-Depression</b>	Geringer Beschwerdedruck	32	9,344	3,721	8,812	5,343					
	Hohe Beschwerdedruck	27	12,259	5,972	10,815	5,061					
	Gesamt	59	10,678	5,053	9,729	5,268	0,013	0,773	0,044	0,081	0,415
<b>MMPI-Hysteria</b>	Geringer Beschwerdedruck	31	9,774	4,877	8,677	5,250					
	Hohe Beschwerdedruck	25	14,720	6,393	13,240	4,781					
	Gesamt	56	11,982	6,080	10,714	5,499	0,166	0,630	0,001	0,039	0,754
<b>MMPI-Psychasthenia</b>	Geringer Beschwerdedruck	32	10,531	4,593	9,500	5,131					
	Hohe Beschwerdedruck	27	13,444	6,191	10,667	5,428					
	Gesamt	59	11,864	5,532	10,034	5,255	0,113	0,764	0,120	0,000	0,096



**Tabelle 74b: Varianzanalyse (Manova): Anfangsbeschwerdedruck zum Zeitpunkt t1, Zusammenhang mit den MMPI(K)-Skalen L, V, C, PD, M/F, PA, SC, HA, SI und VEV im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Anfangsbeschwerdedruck	t3			t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
<b>MMPI-Lie</b>	Geringer Beschwerdedruck	29	3,724	1,730	3,690	1,795					
	Höherer Beschwerdedruck	25	3,400	1,384	3,760	1,393					
	Gesamt	54	3,574	1,573	3,722	1,607	0,262	0,203	0,743	0,437	0,347
<b>MMPI-Validity</b>	Geringer Beschwerdedruck	27	2,630	1,757	2,963	2,084					
	Höherer Beschwerdedruck	22	4,545	3,082	4,545	3,460					
	Gesamt	49	3,490	2,599	3,673	2,868	0,008	0,016	0,016	0,552	0,552
<b>MMPI-Correction</b>	Geringer Beschwerdedruck	31	10,387	3,721	10,355	3,996					
	Höherer Beschwerdedruck	25	9,280	3,835	10,760	3,800					
	Gesamt	56	9,893	3,779	10,536	3,880	0,876	0,795	0,717	0,057	0,047
<b>MMPI-Psychopathic Deviate</b>	Geringer Beschwerdedruck	31	9,290	3,743	8,806	4,045					
	Höherer Beschwerdedruck	26	10,654	4,849	9,923	4,824					
	Gesamt	57	9,912	4,298	9,316	4,412	0,177	0,357	0,252	0,170	0,778
<b>MMPI-Masculinity/Femininity</b>	Geringer Beschwerdedruck	31	14,742	3,444	14,258	3,502					
	Höherer Beschwerdedruck	27	14,593	3,178	14,296	3,688					
	Gesamt	58	14,672	3,295	14,276	3,558	0,672	0,786	0,950	0,081	0,670
<b>MMPI-Paranoia</b>	Geringer Beschwerdedruck	32	5,156	2,529	4,469	2,355					
	Höherer Beschwerdedruck	24	7,583	3,361	7,333	3,795					
	Gesamt	56	6,196	3,130	5,696	3,346	0,146	0,016	0,001	0,109	0,451
<b>MMPI-Schizophrenia</b>	Geringer Beschwerdedruck	26	5,731	2,987	5,692	3,586					
	Höherer Beschwerdedruck	23	9,087	6,179	6,783	3,813					
	Gesamt	49	7,306	5,001	6,204	3,697	0,001	0,766	0,049	0,032	0,038
<b>MMPI-Hypomania</b>	Geringer Beschwerdedruck	31	9,839	3,634	9,290	3,934					
	Höherer Beschwerdedruck	24	10,250	3,904	9,750	3,566					
	Gesamt	55	10,018	3,724	9,491	3,751	0,713	0,613	0,652	0,149	0,946
<b>MMPI-Social Introversion</b>	Geringer Beschwerdedruck	31	12,839	4,879	13,032	5,363					
	Höherer Beschwerdedruck	26	13,577	4,100	11,962	5,000					
	Gesamt	57	13,175	4,516	12,544	5,183	0,363	0,713	0,890	0,159	0,075
<b>VEV</b>	Geringer Beschwerdedruck	31	197,839	29,270	199,774	34,864					
	Höherer Beschwerdedruck	27	194,037	41,006	202,852	51,482					
	Gesamt	58	196,069	34,951	210,207	43,025	0,080	0,043	0,965	0,411	0,598

**Tabelle 75a: Varianzanalyse (Manova): Anfangsbeschwerdedruck zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5 den MMPI(K)-Skalen HS, D, HY, PT im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Anfangsbeschwerdedruck	t1			t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
<b>GBB1-Erschöpfungsneigung</b>	Geringer Beschwerdedruck	34	6,853	3,986	4,559	4,165					
	Hoher Beschwerdedruck	33	15,030	4,606	8,879	5,243					
	Gesamt	67	10,881	5,933	6,687	5,170	0,410	0,191	0,000	0,000	0,004
<b>GBB2-Magenbeschwerden</b>	Geringer Beschwerdedruck	33	3,030	3,618	1,970	2,801					
	Hoher Beschwerdedruck	32	7,750	3,927	4,125	3,563					
	Gesamt	65	5,354	4,435	3,031	3,354	0,646	0,179	0,000	0,000	0,010
<b>GBB3-Gliederschmerzen</b>	Geringer Beschwerdedruck	34	4,853	3,491	3,971	3,647					
	Hoher Beschwerdedruck	33	12,818	3,762	8,000	4,697					
	Gesamt	67	8,776	5,390	5,955	4,633	0,670	0,151	0,000	0,000	0,001
<b>GBB4-Herzbeschwerden</b>	Geringer Beschwerdedruck	34	4,029	2,504	2,235	2,323					
	Hoher Beschwerdedruck	33	7,758	4,176	3,697	4,012					
	Gesamt	67	5,866	3,888	2,955	3,323	0,004	0,002	0,000	0,000	0,038
<b>GBB5-Beschwerdedruck</b>	Geringer Beschwerdedruck	33	18,515	7,942	12,667	10,626					
	Hoher Beschwerdedruck	32	43,125	9,967	25,063	13,174					
	Gesamt	65	30,631	15,277	18,769	13,399	0,204	0,230	0,000	0,000	0,001
<b>MMPI-Hypochondriasis</b>	Geringer Beschwerdedruck	33	6,455	3,032	4,030	3,795					
	Hoher Beschwerdedruck	30	12,833	4,153	8,467	4,681					
	Gesamt	63	9,492	4,809	6,143	4,762	0,085	0,248	0,000	0,000	0,085
<b>MMPI-Depression</b>	Geringer Beschwerdedruck	34	12,412	4,179	8,412	5,433					
	Hoher Beschwerdedruck	32	15,812	5,076	10,250	5,657					
	Gesamt	66	14,061	4,908	9,303	5,577	0,276	0,821	0,014	0,000	0,281
<b>MMPI-Hysteria</b>	Geringer Beschwerdedruck	31	12,613	4,681	8,387	5,110					
	Hoher Beschwerdedruck	30	17,933	5,212	12,400	6,123					
	Gesamt	61	15,230	5,593	10,361	5,939	0,560	0,327	0,000	0,000	0,448
<b>MMPI-Psychasthenia</b>	Geringer Beschwerdedruck	32	13,656	5,103	8,906	5,300					
	Hoher Beschwerdedruck	31	16,710	5,454	10,452	5,494					
	Gesamt	63	15,159	5,457	9,667	5,409	0,713	0,842	0,045	0,000	0,315

**Tabelle 75b: Varianzanalyse (Manova): Anfangsbeschwerdedruck zum Zeitpunkt t1, Zusammenhang mit den MMPI(K)-Skalen L, V, C, PD, M/F, PA, SC, HA, SI im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Anfangsbeschwerdedruck	N	t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
			M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Geringer Beschwerdedruck	32	3,406	2,168	3,594	1,811					
	Höherer Beschwerdedruck	28	3,679	2,001	3,714	1,584					
	Gesamt	60	3,533	2,079	3,650	1,696	0,669	0,474	0,642	0,669	0,771
MMPI-Validity	Geringer Beschwerdedruck	27	3,815	2,321	2,963	2,084					
	Höherer Beschwerdedruck	29	6,724	4,898	4,069	3,206					
	Gesamt	56	5,321	4,113	3,536	2,757	0,000	0,029	0,009	0,001	0,079
MMPI-Correction	Geringer Beschwerdedruck	33	9,242	3,192	10,576	4,000					
	Höherer Beschwerdedruck	29	8,172	3,901	10,207	3,717					
	Gesamt	62	8,742	3,553	10,403	3,843	0,277	0,690	0,364	0,002	0,504
MMPI-Psychopathic Deviate	Geringer Beschwerdedruck	31	11,226	4,201	8,935	3,983					
	Höherer Beschwerdedruck	29	13,586	4,524	10,103	4,821					
	Gesamt	60	12,367	4,483	9,500	4,409	0,693	0,309	0,062	0,000	0,365
MMPI-Masculinity/Femininity	Geringer Beschwerdedruck	31	15,290	3,154	14,516	3,623					
	Höherer Beschwerdedruck	30	14,867	3,857	14,500	3,928					
	Gesamt	61	15,082	3,494	14,508	3,744	0,276	0,660	0,805	0,062	0,499
MMPI-Paranoia	Geringer Beschwerdedruck	30	6,133	2,713	4,667	2,440					
	Höherer Beschwerdedruck	28	9,857	4,428	7,357	3,572					
	Gesamt	58	7,931	4,069	5,966	3,303	0,012	0,048	0,000	0,000	0,316
MMPI-Schizophrenia	Geringer Beschwerdedruck	26	8,077	3,566	5,115	3,570					
	Höherer Beschwerdedruck	27	11,444	6,060	6,667	3,961					
	Gesamt	53	9,792	5,231	5,906	3,819	0,009	0,600	0,010	0,000	0,191
MMPI-Hypomania	Geringer Beschwerdedruck	29	10,069	3,035	9,138	4,129					
	Höherer Beschwerdedruck	29	10,759	3,429	9,724	3,514					
	Gesamt	58	10,414	3,228	9,431	3,812	0,522	0,399	0,447	0,022	0,902
MMPI-Social Introversion	Geringer Beschwerdedruck	32	14,000	4,436	12,750	5,512					
	Höherer Beschwerdedruck	30	14,033	4,131	11,600	5,379					
	Gesamt	62	14,016	4,256	12,194	5,435	0,699	0,894	0,599	0,007	0,375

**Tabelle 76a: Varianzanalyse (Manova): Anfangsdepression zum Zeitpunkt t1, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Anfangsdepression		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
<b>GBB1-Erschöpfungsneigung</b>	Geringe Anfangsdep.	53	8,434	5,002	6,264	4,395					
	Hohe Anfangsdep.	56	13,107	5,308	8,339	5,866					
	Gesamt	109	10,835	5,648	7,330	5,283	0,665	0,036	0,000	0,000	0,012
<b>GBB2-Magenbeschwerden</b>	Geringe Anfangsdep.	53	3,811	3,509	2,698	2,926					
	Hohe Anfangsdep.	56	6,875	4,656	4,411	3,622					
	Gesamt	109	5,385	4,397	3,578	3,397	0,040	0,041	0,000	0,000	0,028
<b>GBB3-Gliederschmerzen</b>	Geringe Anfangsdep.	54	6,630	5,246	4,685	3,414					
	Hohe Anfangsdep.	52	10,173	4,930	6,942	4,478					
	Gesamt	106	8,368	5,372	5,792	4,112	0,656	0,053	0,000	0,000	0,143
<b>GBB4-Herzbeschwerden</b>	Geringe Anfangsdep.	52	4,635	3,320	3,000	3,036					
	Hohe Anfangsdep.	54	6,852	3,579	4,500	4,215					
	Gesamt	106	5,764	3,614	3,764	3,743	0,590	0,020	0,003	0,000	0,298
<b>GBB5-Beschwerdedruck</b>	Geringe Anfangsdep.	51	23,490	11,402	16,627	9,073					
	Hohe Anfangsdep.	48	36,417	13,627	23,625	14,420					
	Gesamt	99	29,758	14,055	20,020	12,413	0,216	0,002	0,000	0,000	0,021
<b>MMPI-Hypochondriasis</b>	Geringe Anfangsdep.	52	6,904	3,291	5,519	3,702					
	Hohe Anfangsdep.	54	12,037	3,981	9,056	4,615					
	Gesamt	106	9,519	4,462	7,321	4,535	0,174	0,115	0,000	0,000	0,038
<b>MMPI-Depression</b>	Geringe Anfangsdep.	54	10,833	3,076	8,444	4,036					
	Hohe Anfangsdep.	54	17,574	2,098	13,333	4,941					
	Gesamt	108	14,204	4,282	10,889	5,118	0,006	0,144	0,000	0,000	0,024
<b>MMPI-Hysteria</b>	Geringe Anfangsdep.	50	12,320	3,930	10,160	5,040					
	Hohe Anfangsdep.	50	18,720	3,985	14,500	5,950					
	Gesamt	100	15,520	5,084	12,330	5,903	0,923	0,249	0,000	0,000	0,043
<b>MMPI-Psychasthenia</b>	Geringe Anfangsdep.	51	11,412	4,016	9,137	4,741					
	Hohe Anfangsdep.	54	18,241	3,425	14,593	5,160					
	Gesamt	105	14,924	5,049	11,943	5,646	0,255	0,544	0,000	0,000	0,142

**Tabelle 76b: Varianzanalyse (Manova): Anfangsdepression zum Zeitpunkt t1, Zusammenhang mit den MMPI(K)-Skalen L, V, C, PD, M/F, PA, SC, HA, SI im Einjahresvergleich von t1 bis t3 in der Untersuchungsgruppe**

	Anfangsdepression	t1			t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Geringere Anfangsdep.	51	3,510	1,701	3,922	1,820	0,242	0,905	0,867	0,134	0,323
	Höhere Anfangsdep.	47	3,617	2,017	3,702	1,852					
	Gesamt	98	3,561	1,850	3,816	1,830					
MMPI-Validity	Geringere Anfangsdep.	44	3,591	2,509	3,295	2,339	0,000	0,002	0,000	0,014	0,101
	Höhere Anfangsdep.	47	6,766	4,478	5,319	3,794					
	Gesamt	91	5,231	3,975	4,341	3,317					
MMPI-Correction	Geringere Anfangsdep.	52	10,423	3,152	11,231	3,462	0,061	0,984	0,000	0,000	0,254
	Höhere Anfangsdep.	54	6,778	2,424	8,278	3,472					
	Gesamt	106	8,566	3,338	9,726	3,756					
MMPI-Psychopathic Deviate	Geringere Anfangsdep.	47	10,277	3,987	8,957	3,689	0,317	0,095	0,000	0,000	0,099
	Höhere Anfangsdep.	52	14,865	3,453	12,288	4,692					
	Gesamt	99	12,687	4,356	10,707	4,543					
MMPI-Masculinity/Femininity	Geringere Anfangsdep.	50	14,840	3,106	14,560	3,284	0,938	0,643	0,464	0,047	0,539
	Höhere Anfangsdep.	53	15,396	3,140	14,868	3,076					
	Gesamt	103	15,126	3,121	14,718	3,167					
MMPI-Paranoia	Geringere Anfangsdep.	48	5,896	2,570	5,333	2,478	0,006	0,000	0,000	0,002	0,163
	Höhere Anfangsdep.	50	9,700	3,861	8,260	4,184					
	Gesamt	98	7,837	3,793	6,827	3,739					
MMPI-Schizophrenia	Geringere Anfangsdep.	45	6,644	3,069	5,267	3,122	0,019	0,000	0,000	0,000	0,014
	Höhere Anfangsdep.	52	12,596	4,326	9,077	5,306					
	Gesamt	97	9,835	4,812	7,309	4,803					
MMPI-Hypomania	Geringere Anfangsdep.	46	9,022	2,671	8,478	3,740	0,895	0,316	0,000	0,076	0,868
	Höhere Anfangsdep.	51	11,549	2,723	11,098	3,233					
	Gesamt	97	10,351	2,969	9,856	3,705					
MMPI-Social Introversion	Geringere Anfangsdep.	52	12,058	4,604	11,673	4,596	0,058	0,934	0,000	0,003	0,055
	Höhere Anfangsdep.	56	15,946	3,539	14,161	4,544					
	Gesamt	108	14,074	4,511	12,963	4,716					

**Tabelle 77a: Varianzanalyse (Manova): Anfangsdepression, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Anfangsdepression		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
GBB1-Erschöpfungsneigung	Geringe Anfangsdep.	33	6,606	3,699	5,909	4,619	0,002	0,361	0,062	0,253	0,793
	Hohe Anfangsdep.	32	8,719	6,517	8,281	5,437					
	Gesamt	65	7,646	5,343	7,077	5,140					
GBB2-Magenbeschwerden	Geringe Anfangsdep.	31	2,935	2,966	2,710	2,866	0,490	0,412	0,208	0,118	0,364
	Hohe Anfangsdep.	32	4,156	3,361	3,313	3,326					
	Gesamt	63	3,556	3,207	3,016	3,098					
GBB3-Gliederschmerzen	Geringe Anfangsdep.	33	4,667	3,159	5,182	3,531	0,015	0,074	0,033	0,570	0,092
	Hohe Anfangsdep.	32	7,469	4,899	6,438	4,866					
	Gesamt	65	6,046	4,314	5,800	4,254					
GBB4-Herzbeschwerden	Geringe Anfangsdep.	33	3,727	3,319	3,030	2,640	0,041	0,401	0,708	0,005	0,254
	Hohe Anfangsdep.	31	4,484	4,830	2,871	3,074					
	Gesamt	64	4,094	4,105	2,953	2,836					
GBB5-Beschwerdedruck	Geringe Anfangsdep.	31	18,000	8,177	16,806	11,089	0,000	0,286	0,074	0,034	0,254
	Hohe Anfangsdep.	31	24,774	16,250	20,871	13,502					
	Gesamt	62	21,387	13,206	18,839	12,423					
MMPI-Hypochondriasis	Geringe Anfangsdep.	32	5,188	3,053	5,094	3,913	0,009	0,462	0,001	0,013	0,023
	Hohe Anfangsdep.	31	9,613	4,958	7,548	4,471					
	Gesamt	63	7,365	4,640	6,302	4,343					
MMPI-Depression	Geringe Anfangsdep.	33	8,515	3,073	8,788	5,030	0,002	0,847	0,003	0,017	0,004
	Hohe Anfangsdep.	31	13,355	5,407	10,516	5,208					
	Gesamt	64	10,859	4,966	9,625	5,150					
MMPI-Hysteria	Geringe Anfangsdep.	33	10,000	4,323	9,606	4,534	0,026	0,104	0,006	0,006	0,035
	Hohe Anfangsdep.	28	14,857	6,542	11,964	6,131					
	Gesamt	61	12,230	5,931	10,689	5,412					
MMPI-Psychasthenia	Geringe Anfangsdep.	32	9,625	4,109	9,500	5,029	0,092	0,526	0,007	0,000	0,000
	Hohe Anfangsdep.	31	14,839	5,586	11,032	5,642					
	Gesamt	63	12,190	5,518	10,254	5,352					

**Tabelle 77b: Varianzanalyse (Manova): Anfangsdepression, Zusammenhang mit den MMPI(K)-Skalen  
L, V, C, PD, M/F, PA, SC, HA, SI und VEV im Zweijahresvergleich von t3 bis t5 in der Untersuchungsgruppe**

	Anfangsdepression		t3		t5		t3	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Geringere Anfangsdep.	31	3,548	1,588	3,516	1,525	0,769	0,711	0,720	0,464	0,371
	Höhere Anfangsdep.	28	3,500	1,503	3,821	1,634					
	Gesamt	59	3,525	1,535	3,661	1,571					
MMPI-Validity	Geringere Anfangsdep.	29	2,655	1,758	2,862	1,747	0,000	0,001	0,002	0,397	0,164
	Höhere Anfangsdep.	25	5,400	4,153	4,560	3,429					
	Gesamt	54	3,926	3,369	3,648	2,769					
MMPI-Correction	Geringere Anfangsdep.	33	10,848	3,633	10,667	3,830	0,751	0,972	0,092	0,050	0,014
	Höhere Anfangsdep.	28	8,464	3,426	10,036	3,805					
	Gesamt	61	9,754	3,709	10,377	3,800					
MMPI-Psychopathic Deviate	Geringere Anfangsdep.	33	8,697	3,687	8,909	3,803	0,281	0,118	0,048	0,112	0,036
	Höhere Anfangsdep.	28	11,607	4,500	10,107	5,080					
	Gesamt	61	10,033	4,301	9,459	4,437					
MMPI-Masculinity/Femininity	Geringere Anfangsdep.	33	14,424	3,536	13,545	3,709	0,295	0,270	0,152	0,324	0,002
	Höhere Anfangsdep.	60	14,933	2,924	15,400	3,035					
	Gesamt	63	14,667	3,243	14,429	3,504					
MMPI-Paranoia	Geringere Anfangsdep.	32	5,375	2,240	4,906	2,205	0,001	0,002	0,006	0,012	0,258
	Höhere Anfangsdep.	29	7,931	4,276	6,724	3,918					
	Gesamt	61	6,590	3,575	5,770	3,242					
MMPI-Schizophrenia	Geringere Anfangsdep.	27	5,407	2,791	5,556	3,226	0,001	0,238	0,005	0,019	0,009
	Höhere Anfangsdep.	26	9,692	5,641	7,192	4,079					
	Gesamt	53	7,509	4,886	6,358	3,727					
MMPI-Hypomania	Geringere Anfangsdep.	32	9,031	3,772	8,719	3,419	0,348	0,650	0,021	0,208	0,723
	Höhere Anfangsdep.	27	11,185	3,163	10,630	3,723					
	Gesamt	59	10,017	3,641	9,593	3,658					
MMPI-Social Introversion	Geringere Anfangsdep.	31	12,097	4,407	12,355	5,559	0,883	0,738	0,359	0,197	0,077
	Höhere Anfangsdep.	29	14,103	4,287	12,483	5,221					
	Gesamt	60	13,067	4,430	12,417	5,353					
VEV	Geringere Anfangsdep.	31	198,774	28,621	199,677	34,980	0,069	0,050	0,757	0,240	0,302
	Höhere Anfangsdep.	31	194,839	40,138	208,516	50,414					
	Gesamt	62	196,806	34,629	204,097	43,262					

**Tabelle 78a: Varianzanalyse (Manova): Anfangsdepression, Zusammenhang mit den GBB-Skalen 1 bis 5, den MMPI(K)-Skalen HS, D, HY, PT im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Anfangsdepression		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
GBB1-Erschöpfungsneigung	Geringe Anfangsdep.	38	8,289	4,608	5,447	4,524	0,168	0,344	0,000	0,000	0,013
	Hohe Anfangsdep.	35	14,086	5,813	8,086	5,305					
	Gesamt	73	11,068	5,947	6,712	5,057					
GBB2-Magenbeschwerden	Geringe Anfangsdep.	35	4,057	3,850	2,657	2,828	0,344	0,128	0,015	0,000	0,055
	Hohe Anfangsdep.	34	6,912	4,535	3,647	3,684					
	Gesamt	69	5,464	4,411	3,145	3,291					
GBB3-Gliederschmerzen	Geringe Anfangsdep.	38	7,158	5,139	5,053	3,448	0,908	0,007	0,007	0,000	0,103
	Hohe Anfangsdep.	33	10,848	5,038	6,758	5,494					
	Gesamt	71	8,873	5,385	5,845	4,563					
GBB4-Herzbeschwerden	Geringe Anfangsdep.	36	5,028	3,939	3,000	2,694	0,777	0,020	0,082	0,000	0,079
	Hohe Anfangsdep.	35	7,171	3,753	3,286	4,019					
	Gesamt	71	6,085	3,970	3,141	3,390					
GBB5-Beschwerdedruck	Geringe Anfangsdep.	33	22,879	11,349	16,061	10,954	0,139	0,069	0,000	0,000	0,004
	Hohe Anfangsdep.	32	38,625	14,797	21,562	15,195					
	Gesamt	65	30,631	15,277	18,769	13,399					
MMPI-Hypochondriasis	Geringe Anfangsdep.	36	6,694	3,152	4,778	3,765	0,108	0,095	0,000	0,000	0,002
	Hohe Anfangsdep.	33	12,758	4,169	7,667	5,035					
	Gesamt	69	9,594	4,754	6,159	4,620					
MMPI-Depression	Geringe Anfangsdep.	37	10,351	3,352	8,081	5,187	0,020	0,600	0,000	0,000	0,000
	Hohe Anfangsdep.	35	18,200	2,247	10,229	5,673					
	Gesamt	72	14,167	4,870	9,125	5,498					
MMPI-Hysteria	Geringe Anfangsdep.	34	11,676	4,110	9,088	4,522	0,728	0,026	0,000	0,000	0,001
	Hohe Anfangsdep.	33	19,273	3,867	11,576	6,713					
	Gesamt	67	15,418	5,508	10,313	5,800					
MMPI-Psychasthenia	Geringe Anfangsdep.	35	11,429	4,168	8,543	5,181	0,157	0,525	0,000	0,000	0,000
	Hohe Anfangsdep.	33	19,364	3,248	10,758	5,793					
	Gesamt	68	15,279	5,460	9,618	5,558					



**Tabelle 78b: Varianzanalyse (Manova): Anfangsdepression zum Zeitpunkt t1, Zusammenhang mit den MMPI(K)-Skalen L, V, C, PD, M/F, PA, SC, HA, SI im Dreijahresvergleich von t1 bis t5 in der Untersuchungsgruppe**

	Anfangsdepression		t1		t5		t1	t5	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	COCH P	COCH P	F	F	F
MMPI-Lie	Geringere Anfangsdep.	36	3,389	2,018	3,528	1,665	0,990	0,779	0,486	0,579	0,991
	Höhere Anfangsdep.	30	3,667	2,023	3,800	1,750					
	Gesamt	66	3,515	2,010	3,652	1,696					
MMPI-Validity	Geringere Anfangsdep.	31	3,677	2,414	2,903	1,680	0,000	0,000	0,000	0,000	0,006
	Höhere Anfangsdep.	30	7,667	4,737	4,167	3,323					
	Gesamt	61	5,639	4,219	3,525	2,675					
MMPI-Correction	Geringere Anfangsdep.	36	10,028	3,452	10,750	3,722	0,223	0,809	0,009	0,001	0,039
	Höhere Anfangsdep.	31	7,097	2,785	9,839	3,882					
	Gesamt	67	8,672	3,466	10,328	3,795					
MMPI-Psychopathic Deviate	Geringere Anfangsdep.	31	10,258	4,123	8,968	4,070	0,378	0,315	0,002	0,000	0,007
	Höhere Anfangsdep.	34	14,676	3,522	10,029	4,871					
	Gesamt	65	12,569	4,394	9,523	4,504					
MMPI-Masculinity/Femininity	Geringere Anfangsdep.	33	14,848	3,447	13,818	3,877	0,966	0,432	0,179	0,069	0,069
	Höhere Anfangsdep.	32	15,469	3,473	15,469	3,370					
	Gesamt	65	15,154	3,447	14,631	3,702					
MMPI-Paranoia	Geringere Anfangsdep.	31	6,032	2,627	5,129	2,172	0,009	0,003	0,000	0,000	0,007
	Höhere Anfangsdep.	33	10,273	4,245	6,758	3,792					
	Gesamt	64	8,219	4,123	5,969	3,197					
MMPI-Schizophrenia	Geringere Anfangsdep.	30	6,267	3,107	4,767	3,115	0,158	0,123	0,000	0,000	0,000
	Höhere Anfangsdep.	27	13,741	4,072	7,185	4,188					
	Gesamt	57	9,807	5,184	5,912	3,828					
MMPI-Hypomania	Geringere Anfangsdep.	33	9,303	2,687	8,424	3,640	0,297	0,915	0,004	0,026	0,915
	Höhere Anfangsdep.	29	11,586	3,257	10,621	3,570					
	Gesamt	62	10,371	3,158	9,452	3,745					
MMPI-Social Introversion	Geringere Anfangsdep.	35	12,086	4,083	11,886	5,624	0,152	0,896	0,017	0,001	0,002
	Höhere Anfangsdep.	32	16,438	3,172	12,313	5,497					
	Gesamt	67	14,164	4,255	12,090	5,526					

**Tabelle 79a: Varianzanalyse (Manova): Untersuchungsgruppe (UG) und Kontrollgruppe (KG) im Einjahresvergleich von Messzeitpunkt t1 bis Messzeitpunkt t3, geprüft in den fünf GBB-Skalen und den MMPI(K)-Skalen HS, D, HY, PT**

	UG/KG		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	BARTLETT P	BARTLETT P	F	F	F
GBB1-Erschöpfungsneigung	Untersuchungsgruppe	100	11,180	5,620	7,940	5,079					
	Kontrollgruppe	37	10,297	5,542	9,649	6,533					
	Gesamt	137	10,942	5,592	8,401	5,537	0,920	0,059	0,659	0,000	0,014
GBB2-Magenbeschwerden	Untersuchungsgruppe	86	6,488	4,183	4,453	3,296					
	Kontrollgruppe	37	5,351	3,713	5,297	4,129					
	Gesamt	123	6,146	4,066	4,707	3,571	0,407	0,102	0,830	0,000	0,002
GBB3-Gliederschmerzen	Untersuchungsgruppe	96	9,010	5,172	6,375	3,878					
	Kontrollgruppe	38	9,658	5,851	8,474	6,030					
	Gesamt	134	9,194	5,359	6,970	4,665	0,362	0,001	0,108	0,000	0,103
GBB4-Herzbeschwerden	Untersuchungsgruppe	87	6,391	3,509	4,563	3,669					
	Kontrollgruppe	38	6,816	4,643	5,447	5,505					
	Gesamt	125	5,222	3,874	4,832	4,306	0,039	0,002	0,364	0,000	0,507
GBB5-Beschwerdedruck	Untersuchungsgruppe	99	29,758	14,055	20,020	12,413					
	Kontrollgruppe	37	32,000	16,546	28,946	18,652					
	Gesamt	136	30,368	14,745	22,449	14,850	0,228	0,002	0,028	0,000	0,008
MMPI-Hypochondriasis	Untersuchungsgruppe	102	9,814	4,230	7,529	4,487					
	Kontrollgruppe	38	9,737	4,578	9,079	5,385					
	Gesamt	140	9,793	4,311	7,950	4,778	0,559	0,171	0,350	0,000	0,024
MMPI-Depression	Untersuchungsgruppe	108	14,204	4,282	10,889	5,118					
	Kontrollgruppe	37	13,027	5,069	12,027	6,130					
	Gesamt	145	13,903	4,507	11,179	5,395	0,206	0,175	0,982	0,000	0,004
MMPI-Hysteria	Untersuchungsgruppe	100	15,520	5,084	12,330	5,903					
	Kontrollgruppe	37	15,649	5,638	14,108	6,790					
	Gesamt	137	15,555	5,219	12,810	6,181	0,448	0,302	0,339	0,000	0,083
MMPI-Psychasthenia	Untersuchungsgruppe	104	15,000	5,013	12,058	5,549					
	Kontrollgruppe	34	12,735	5,997	12,441	6,574					
	Gesamt	138	14,442	5,340	12,152	5,796	0,195	0,220	0,348	0,000	0,004

**Tabelle 79b: Varianzanalyse (Manova): Untersuchungsgruppe (UG) und Kontrollgruppe (KG) im Einjahresvergleich von Messzeitpunkt t1 bis Messzeitpunkt t3, geprüft in den MMPI(K)-Skalen L, V, C, PD, M/F, PA, SC, HA, SI**

	UG/KG		t1		t3		t1	t3	Signifikanz des Unterschiedes zwischen Subgruppen	Signifikanz des Zeiteffektes	Signifikanz des Unterschiedes im Zeitverlauf
		N	M	s	M	s	BARTLETT P	BARTLETT P	F	F	F
MMPI-Lie	Untersuchungsgruppe	93	3,753	1,698	3,946	1,784	0,368	0,815	0,797	0,297	0,022
	Kontrollgruppe	37	4,027	1,922	3,514	1,726					
	Gesamt	130	3,831	1,762	3,823	1,772					
MMPI-Validity	Untersuchungsgruppe	85	5,600	3,852	4,482	3,358	0,724	0,574	0,686	0,037	0,267
	Kontrollgruppe	29	4,931	3,644	4,586	3,660					
	Gesamt	114	5,430	3,796	4,509	3,421					
MMPI-Correction	Untersuchungsgruppe	106	8,566	3,338	9,726	3,756	0,841	0,760	0,269	0,011	0,212
	Kontrollgruppe	35	8,257	3,433	8,657	3,597					
	Gesamt	141	8,489	3,352	9,461	3,733					
MMPI-Psychopathic Deviate	Untersuchungsgruppe	99	12,687	4,356	10,707	4,543	0,134	0,046	0,272	0,009	0,031
	Kontrollgruppe	31	10,806	5,400	10,613	6,026					
	Gesamt	130	12,238	4,674	10,685	4,912					
MMPI-Masculinity/Femininity	Untersuchungsgruppe	103	15,126	3,121	14,718	3,167	0,780	0,109	0,029	0,159	0,617
	Kontrollgruppe	36	13,722	3,001	13,528	3,931					
	Gesamt	139	14,763	3,141	14,410	3,407					
MMPI-Paranoia	Untersuchungsgruppe	98	7,837	3,793	6,827	3,739	0,510	0,446	0,971	0,013	0,521
	Kontrollgruppe	35	7,657	4,158	7,057	4,158					
	Gesamt	133	7,789	3,877	6,887	3,839					
MMPI-Schizophrenia	Untersuchungsgruppe	97	9,835	4,312	7,309	4,803	0,660	0,621	0,493	0,011	0,003
	Kontrollgruppe	31	9,097	5,134	9,290	5,165					
	Gesamt	128	9,656	4,882	7,789	4,946					
MMPI-Hypomania	Untersuchungsgruppe	97	10,351	2,969	9,856	3,705	0,192	0,082	0,192	0,174	0,790
	Kontrollgruppe	30	11,167	3,592	10,833	4,764					
	Gesamt	127	10,543	3,131	10,087	3,982					
MMPI-Social Introversion	Untersuchungsgruppe	108	14,074	4,511	12,963	4,716	0,064	0,126	0,266	0,025	0,322
	Kontrollgruppe	37	12,757	5,761	12,324	5,779					
	Gesamt	145	13,738	4,873	12,800	4,995					

**Tabelle 83: Die Teilnehmer an der Katamnese  
nach soziodemografischen Merkmalen  
(zum Zeitpunkt der Katamneseerhebung)**

**Tabelle: 83a**

**Katamneseteilnehmer: Geschlecht**

	<b>N</b>	<b>Prozent</b>
Männer	16	28,1
Frauen	41	71,9
Gesamt	57	100,0

**Tabelle 83b**

**Katamneseteilnehmer: Altersgruppe**

		<b>Häufigkeit</b>	<b>Prozent</b>	<b>Gültige Prozente</b>	<b>Kumulierte Prozente</b>
gültig	21-30	1	1,8	1,8	1,8
	31-40	27	47,4	48,2	50,0
	41-50	24	42,1	42,9	92,9
	51-60	4	7,0	7,1	100,0
	Gesamt	56	98,2	100,0	
fehlend		1	1,8		
Gesamt		57	100,0		

**Tabelle 83c**  
**Katamneseteilnehmer: Familienstand**

	Häufigkeit	Prozent	Kumulierte Prozente
ledig	18	31,6	31,6
verheiratet	27	47,4	78,9
getrennt lebend	4	7,0	86,0
geschieden	8	14,0	100,0
Gesamt	57	100,0	

**Tabelle 83d**  
**Katamneseteilnehmer: Anzahl der Kinder**

		N	Prozent	Gültige Prozente
gültig	0	23	40,4	42,6
	1	8	14,0	14,8
	2	17	29,8	31,5
	3	5	8,8	9,3
	4	1	1,8	1,9
	Gesamt	54	94,7	100,0
fehlend		3	5,3	
Gesamt		57	100,0	

**Tabelle 83e****Katamneseteilnehmer: Schulabschluss**

		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
gültig	Hauptschule	4	7,0	7,4	7,4
	Mittlere Reife	15	26,3	27,8	35,2
	Abitur	35	61,4	64,8	100,0
	Gesamt	54	94,7	100,0	
fehlend		3	5,3		
Gesamt		57	100,0		

**Tabelle 83f****Katamneseteilnehmer: Berufsklassifikation**

	Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Freier Beruf, Leiter von Unternehmen	5	8,8	8,8	8,8
Höherer Beamter, leitd. Angestellter	5	8,8	8,8	17,5
Beamter, Angestellter	32	56,1	56,1	73,7
Selbst. Gewerbe- treibender, Landwirt	6	10,5	10,5	84,2
Facharbeiter	1	1,8	1,8	86,0
Hausfrau/-mann, mithelfender Familienangehöriger	3	5,3	5,3	91,2
Schüler, Azubi, Student	2	3,5	3,5	94,7
Rentner, Pensionär	1	1,8	1,8	96,5
arbeitslos	2	3,5	3,5	100,0
Gesamt	57	100,0	100,0	

**2. Auswertung der Katamnesefrage: „Schildern Sie bitte möglichst konkret eine freudige Situation aus der Therapie.“**

### **Auswertung der Frage: „Schildern Sie bitte möglichst konkret eine freudige Situation aus der Therapie“:**

Wir erhielten insgesamt 62 Katamnesefragebögen ausgefüllt zurück. In 39 Fällen war die o.g. Frage nach der Freude beantwortet, in 34 Fällen war die Frage nicht beantwortet.

Die Art der Stellungnahmen war sehr unterschiedlich. In der überwiegenden Mehrzahl der Fälle wurden allgemeine Therapiebewertungen im Sinne von „Ich habe mich gefreut, dass/als es mir besser ging“ abgegeben. Am zweithäufigsten wurden prozessorientierte Darstellungen geliefert, die einen Einblick in die therapeutische Beziehung geben. Teilweise wurden auch einzelne freudige Therapiesituationen geschildert, die jedoch isoliert wiedergegeben werden und dadurch keinen weiterführenden Blick auf den Gesamtprozess freigeben. Einige wenige Äußerungen sind ausweichend, abwehrend und eher nichtssagend. In zwei Fällen wurde die Frage nach der Freude zum Anlass genommen, Negativäußerungen und Beschwerden gegenüber dem therapeutischen Geschehen insgesamt abzugeben.

Hier die Antworten im Einzelnen (um einen möglichst getreuen Einblick zu gewähren hier wortwörtlich wiedergegeben, so wie die Patienten antworteten):

- 1) F301151: Nach jahrelanger Arbeit ist die Freude darüber riesengroß, dass es mir nach einigen Monaten viel, viel besser geht.
- 2) D240849: Ich hecke gerne etwas aus, wie Max und Moritz. In einer Therapiestunde durfte ich Moritz spielen und meine Therapeutin in dieses Spiel mit einbeziehen. Streiche zu spielen, natürlich ohne dabei jemanden zu verletzen, hat einfach Spaß gemacht.
- 3) W250162: Als ich meinen jetzigen Mann kennenlernte und dabei in einer Sitzung unser langsames Kennenlernen schilderte.
- 4) Z130547: Als ich von der Couch auf den Sessel gewechselt bin.
- 5) M160163: Situation: Als es mir sehr schlecht ging, und ich nicht weiter wusste, erzählte er mir, dass es vielen Menschen so geht „und ich weinte“ und es brach vieles aus mir raus. Ich muss sagen, bei vielen schlechten Tagen hat er mir sehr geholfen. Ich bin auch sehr traurig, dass die Therapie beendet ist. Ich hoffe, dass die Antworten Ihnen helfen. Mit freundlichem Gruß.
- 6) A131166: „Nachdem ich 20 Jahre nicht beim Zahnarzt war, hatte der Therapeut mit mehr Druck als jemals zuvor erreicht, dass ich zum Zahnarzt gegangen bin und die Behandlung fortgesetzt habe. In der „Drucksituation“ vorher war mir der Therapeut sehr fremd geworden und nach den ersten Zahnarztbehandlungen konnte ich auch nicht sagen, dass ich froh war, das geschafft zu haben. Viele Monate später habe ich die „Aktion“ reflektierend gesagt, dass ich mich durch den Gang zum Zahnarzt dem Leben einer erwachsenen Frau, die gut für sich sorgen kann, etwas näher fühle. Bis auf Weisheitszähne ziehen lassen, habe ich bis heute die Behandlung beim Zahnarzt fortgesetzt.
- 7) D080361: Wenn ich eine Situation neu gesehen und erkannt habe und ich die Situation auch verstanden habe. Am meisten, wenn ich es dann auch umsetzen kann. Noch besser, wenn es dann auch „HÄLT“.
- 8) R050159: Herr Frings war für mich ein sehr guter Therapeut. Ich hatte immer das Gefühl, dass er jetzt ganz für mich da ist, mir zuhört, mich ernst nimmt und versteht. Darüber war ich sehr dankbar und froh. Es ist nicht selbstverständlich; ich habe leider auch andere Erfahrungen gemacht. Doch manchmal gab es Situationen, die mich/uns besonders berührten. Eine besondere Situation für mich war: Ich hatte einen Termin abgesagt, zu dem ich dann doch gekommen war. Er antwortete: „Nehmen Sie Platz“ in einem warmen Ton und mit einer einladenden Handbewegung. Ich setzte mich hin und hatte das Gefühl zu fallen, zu fallen, fallen und ganz weich in einem Berg von Kissen zu landen. Erwartet hatte ich ein Donnerwetter.
- 9) B141163: Keine konkrete Erinnerung.
- 10) M060364: Situationen, in denen die Therapeutin mein Verhalten, meine Einstellungen überzeichnet oder ironisiert hat, so dass auch ich eine lockere Distanz gewinnen konnte. Das einzige konkrete Beispiel, das ich noch genau im Kopf habe, ist mir zu privat. Der Kommentar der Therapeutin war: „Schön blöd.“



11) W030165: Meine Therapeutin fing gegen Ende der Sitzung häufig an, mit den Fingernägeln zu knipsen, was ich als ein Zeichen von Desinteresse interpretierte und worüber ich mich ziemlich ärgerte. Dass sie nicht an die Decke ging, als ich ihr den folgenden Comic-Strip überreichte, hat mich gefreut. Noch heute freut mich, dass sie mich so gelassen hat, wie ich bin und mich nie von ihr abhängig fühlen musste. Im übrigen ist Therapie harte Arbeit für mich gewesen und meine „Freude“ war immer eher Erleichterung. Gemaltes Comic.

12) K140567: Ich kann mich leider an keine solche Situation erinnern. Ganz im Gegenteil. Ich hatte nie das Gefühl, dass meine Therapeutin auf meine Bedürfnisse die Therapieform zuschnitt. Ich vermisste, dass mir die Therapeutin auch mal Fragen stellte und auch mal die Stunde ankurbelte. Ich fühlte mich stets wie ein Alleinunterhalter, der Monologe zu führen hatte, was nun gar nicht meinem Naturell entspricht. Dies bedingte, dass ich mit immer ablehnenderen Gefühlen und nachher sogar Bauchschmerzen und Unbehagen die Stunde besuchte. Zwei Tage vorher überlegte ich schon, was ich meiner Therapeutin erzählen könnte. Sehr störend empfand ich auch die Einstellung meiner Therapeutin zu Urlaub: Ich durfte nur Urlaub in der Ferien machen, was für mich sehr ungünstig war. Als wenn Urlaub nicht auch zur Erholung und Selbstfindung dient. Alles in allem war ich froh, als die Therapie beendet war. Für mich war es ein Abbruch der Therapie, meine Therapeutin meinte hingegen, die Therapie sei mit positiver Richtung beendet – sehr merkwürdig für mein Empfinden.

13) M190164: Als die Therapiesitzungen insgesamt sich dem Ende näherten und ich nicht mehr auf eine Verlängerung der Therapie gedrängt habe; ich habe das Gefühl genossen, kaum noch Angst vor der Zeit nach der Therapie zu haben. Ich hatte das Gefühl, nun ohne therapeutische Hilfe mit kritischen Situationen und Phasen umgehen zu können und ich habe dies bis heute auch geschafft. (15.02.2000).

14) W020565: Ganz simpel gesagt: Der Moment, in dem ich eine Türklinke anfassen konnte und anschließend nicht mehr dem Zwang des Händewaschens nachkommen musste!

15) W060662: Freude war nicht gerade das Hauptmerkmal der Therapie, da sie meist schmerzhaft war, aber: als ich sagte, ich fühle mich wie Jeanne d'Arc, bemerkte mein Therapeut, in der Zeit von Maschinengewehren und Kalaschnikows sei es vielleicht nicht ganz zeitgemäß, mit Pfeil und Bogen zu schießen.

16) B280740: Ich bin als Kind von sechs Jahren mehrfach vergewaltigt worden. Einmal hätte ich mir verzeihen können, da ich nicht wusste, was mir geschah. Meinem Therapeuten ist es gelungen, mir diese Schuld u. das Gefühl, nicht o.k. zu sein, zu nehmen. Das hat mich befreit und mit tiefer Freude erfüllt.

17) K050967: Ich kann mich an keine konkrete Situation erinnern.

18) B260464: Ich habe gemerkt, dass sich, als ich am Tiefpunkt war, Freunde als Freunde erwiesen haben und Bekannte zu engen Freunden wurden. Mein jüngerer Bruder ist seit dem einer meiner engsten Freunde.

19) M130562: Trotz intensivem Nachdenken ist mir leider keine Situation eingefallen.

20) N010658: Ich kann mich leider an keine konkrete Situation erinnern. Lachen musste ich aber zumeist in Situationen, in denen mir deutlich wurde, dass Motive meines Handelns weniger in „edlen Absichten“, sondern in meinen Bösarbeiten ihren Ursprung haben.

21) H221044: Keine Probleme auf beiden Seiten, als ich eine Stunde geschlafen hatte.

22) S110457: Ein Traum: Fliegen in herrlichem Wetter über meinem Elternhaus und näherer Umgebung, Landung, kurzer Kontakt mit einem kleinen Kind, das mir (?) über die Wange streicht. (Das „innere Kind“?) Dieser Traum berührte mich tief und dann: Frieden und Freude.

23) M020857: Gemeinsames Erkennen, dass eine Situation, durch einen anderen Blickwinkel betrachtet, nicht so ausweglos war wie sie schien.

24) L170359: Als mir gegen Ende der Therapie die Therapeutin meine in den letzten sechs Jahren von mir gemalten Bilder vorlegte. Das erste Bild war ganz erbärmlich, ich war im nachhinein erschrocken, gerührt, „das stimmt ja gar nicht mehr“. Ich kann über manche damaligen Einstellungen und Ansichten lachen. Fühle mich befreit, dankbar dafür, dass ich den Weg raus mit dieser Begleitung finden konnte.

Ich male heute Freudenbilder in kräftigen Farben. Ohne Therapie wäre ich gestorben oder wahnsinnig geworden.

25) S100163: Ausgangssituation = die Verantwortung für meine kranke Mutter lastet auf mir; Übung in der Therapiestunde = Meine Therapeutin steht seitlich von mir, legt mir die Hand auf die Brust, die andere auf den oberen Rücken; ich heule und heule, ich fühle mich gehalten; sie streicht mein Haar und ich fühle mich wunderbar getröstet; alle Verantwortung fällt von mir (ist das Freude?).

26) N130958: Ich habe unter einer starken Antriebsarmut gelitten und dann haben wir uns beide befreit, als ich dann mich unter großer Anstrengung zum Sportkursus angemeldet habe. So war das auch mit vielen anderen Sachen. Ich traute mir kaum etwas zu.

27) G310765: Habe für solche Beschreibung noch nicht genug Abstand zur Therapie, da ich mich noch in Behandlung befinde.

28) K200952: Mein Therapeut hat, glaube ich, Humor. Das war wichtig, weil ich zum Schwarzsehen neige. Ich habe so sicherlich Zuversicht erfahren. Am meisten gefreut habe ich mich, als ich nach den vielen Stunden wirklich begriffen habe, als ich fühlen konnte, dass ich keine Verantwortung mehr für meinen Mann übernehmen muss. Ich habe mich ganz leicht und frei gefühlt. Das war gut. Ich konnte sehen, wie mein Therapeut sich mit mir freut.

29) K280151: Von mir hergestellte Kunstobjekte brachte ich mit in die Therapie. Die Freude (echtes Interesse/Berührtsein) der Therapeutin daran sprang auf mich über und führte zu einer fruchtbaren kunsttherapeutischen Erarbeitung. Ich spürte, dass meine Therapeutin von diesen Momenten profitierte und ich bewunderte ihre an sich fachfremde Kompetenz in diesem Bereich. In diesen Stunden habe ich mich jeweils sehr gefreut und bestärkt gefühlt, meinen kreativen Weg weiter zu verfolgen.

30) K130572: Die freudige Situation besteht darin, dass wenigstens einer einen versteht.

31) K090256: Ich sprach mit meiner Therapeutin sehr oft über meine Familie und konnte durch dieses Sprechen mit ihr viele Dinge hinterher sehr viel positiver sehen.

32) N100156: In einer Therapiestunde ging es um die Bewerbung auf eine befristete (längerfristig) Arbeitsstelle. Durch Studium und Arbeit nach Beendigung meines Studiums habe ich mehr als 16 Jahre nur aushilfsweise gearbeitet (auch während des Studiums), also keine festen oder länger befristeten Arbeitsstellen ergriffen, aus Angst vor zu großer Abhängigkeit. Mein Therapeut erinnerte mich an meine positiven Erfolge in bezug auf andere Abhängigkeiten, hob meine positiven Fähigkeiten hervor, leistete mir Hilfestellung und machte Mut. Ich verlies die Stunde beschwingt, positiv eingestellt und überzeugt, das Richtige zu tun. Ich bekam die Stelle.

33) R160552: Freude darüber, wenn nach längerem Suchen in den Sachverhalten, die Sicht auf das Problem klar wurde. Ich habe mich dann gefreut, einen gangbaren Lösungsweg zu sehen.

34) W130760: In meiner Arbeitsstelle etwas für mich Unangenehmes auszudrücken und dadurch (?) Position geändert zu haben, anstatt nichts zu sagen und nichts zu verändern.

35) W220766: Es war für mich (und meine Therapeutin) oft ein Grund zur Freude, wenn ich wieder einen Schritt weiter war. Sei es ein Satz meiner Therapeutin oder eine Anregung. Ich habe mich auch so gut wie immer auf die Stunde gefreut.

36) E271254: Ich habe mich verliebt und mich in meiner Art und Weise frei und selbstbestimmt mit dieser Liebe gefühlt wie nie zuvor. Der Kommentar der Therapeutin war „wie schön“, aus ganz reinem Herzen und ich habe mich gefreut, so fühlen zu können und bestätigt zu werden.

37) B190951: Die Schilderung einer unerfreulichen Situation fiel mir leichter!

38) R290552: Weiß es nicht mehr genau.

39) W180160: Ich habe gelernt, mich konsequenter zu verhalten. Den anderen deutlich Grenzen zu setzen. Auch lernte ich, gelassener mit Situationen umzugehen.

**3. Erstes Anschreiben an die 230 Psychoanalytiker (DGIP) vom 15.03.1993 und Schreiben des Vorsitzenden der DGPT, Prof. Dr. Streeck, vom 18.09.1992**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. 0221/7408887

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleussner  
Dipl.-Pädagogin Monika Weber

Liebe Kollegin,  
lieber Kollege

wir wenden uns mit der Bitte an Sie, uns bei einem dringenden Forschungsvorhaben zu unterstützen, das u. a. auch Ihre ureigensten Interessen berührt.

## Was Sie wissen sollten:

Seit einiger Zeit häufen sich die Angriffe auf die Psychoanalyse, was deren therapeutische Wirksamkeit anbelangt. Beispielhaft verweisen wir auf Herrn Grawe (Bern), dessen Einschätzung in das Gutachten zum Psychotherapeutengesetz Eingang gefunden hat. Die der DGPT angeschlossenen Gesellschaften wurden kürzlich vom Vorsitzenden der DGPT auf den Ernst der Lage brieflich hingewiesen (siehe Anlage). Individualpsychologen treffen die Aussagen von Herrn Grawe insofern besonders hart, als wir, ginge es nach Herrn Grawe, bis auf weiteres von der Versorgung und Ausbildung auszuschließen wären, da ein Wirksamkeitsnachweis unserer IP-Analysen nicht vorliegt.

## Was wir tun können:

Um nicht Gefahr zu laufen, berufliche Einschränkungen hinnehmen zu müssen, brauchen wir schnellstmöglichst eine empirische Studie, die auch unter Beachtung gängiger wissenschaftstheoretischer Standards die Wirksamkeit unserer Therapien überprüft.

Solch eine Untersuchung durchzuführen und vorzulegen haben wir uns vorgenommen. Wir, das sind Monika Weber (Kinder- und Jugendlichen-Therapeutin DGIP in Practicando-Status), Dieter Schleussner und Franz-Josef Klar (beide Psychoanalytiker DGIP). Als wissenschaftliche Begleiter fungieren Prof. Dr. Günter Heisterkamp und Dr. Rolf Dellen (beide Psychoanalytiker DGIP).

Von der DGIP und dem Düsseldorfer Alfred Adler Institut wurden uns dankenswerterweise bereits Forschungsmittel gewährt.

## Wie Sie uns helfen können:

Was wir jetzt noch benötigen, ist die Unterstützung aller Analytiker DGIP in ganz Deutschland, - folglich auch Ihre Mitarbeit.

Diese sähe so aus, daß Sie sich bereit erklären,

- a) einigen Ihrer Patienten/innen vor, nach einem halben und nach einem Jahr der analytischen Einzeltherapie einen standardisierten Fragebogen (Test) in einem verschlossenen Kuvert zu übergeben und auf die Rückleitung des ausgefüllten Fragebogens an uns zu achten (Bildung der Untersuchungsgruppe).

- b) einem/einigen wartenden Therapiebewerbern von Ihrer Warteliste ebenfalls zur Mitarbeit zu gewinnen, indem Sie ihnen als Gegenleistung für eine aus Forschungsgründen notwendige Wartezeit von ca. 1 Jahr einen festen Therapieplatz zusichern (Bildung der Kontrollgruppe), und ihnen ebenfalls zu Beginn, nach einem halben und nach einem Jahr der Wartezeit den gleichen Fragebogen vorlegen.
- c) den in die Untersuchung eingegangenen Patienten (Untersuchungsgruppe) in Abständen von jeweils einem Jahr und spätestens noch einmal ein Jahr nach Therapieende den gleichen Fragebogen vorzulegen.

Auf Sie kommt dabei keine große Arbeit zu. Ihre Mitarbeit wird sich im wesentlichen darauf beschränken, zwischen den Patienten/innen und uns eine Mittlerrolle zu übernehmen: Sie händigen die Fragebögen aus, nehmen sie zurück und leiten sie an uns weiter.

Wie es weitergeht:

Dieses Schreiben soll zunächst dazu dienen, Sie zur Mitarbeit zu gewinnen. Wenn Sie Ihr Einverständnis erklären, erhalten sie in einem zweiten Schreiben von uns natürlich noch weitere Informationen z. B. zum Versuchsaufbau, zur Art der angewandten Tests sowie Instruktionen für die Probanden. Die Untersuchung soll dann möglichst bald beginnen.

Zusicherungen:

Wir sichern den Patienten einen absolut vertraulichen Umgang mit den Fragebögen zu. Die Daten der Patienten sollen gemäß KV - Chiffrierung anonymisiert an uns weitergeleitet werden.

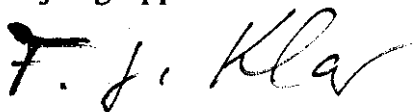
Allerdings benötigen wir für die Dauer der Untersuchung die Zuordnung von Chiffrierung und Praxis, um die Re-Test-Fragebögen gezielt zusenden zu können. Wir möchten Ihnen als Kollegen zusichern, daß wir die Fragebogenergebnisse nicht dazu verwenden werden, Ihre therapeutische Qualifikation zu überprüfen. Da nur ein kleiner, nicht repräsentativer Bruchteil Ihres Patientenstammes in die Untersuchung eingeht, wäre dies auch gar nicht möglich.

Wir hoffen, Sie mit diesem Schreiben hinreichend informiert zu haben, so daß Sie bereit sind, uns mit beiliegendem Antwortcoupon Ihre Unterstützung kurzfristig zuzusichern.

Senden Sie bitte den Antwortcoupon auf jeden Fall bis Anfang April mit dem anliegenden Freiumschlag an uns zurück.

Für Ihre Bemühungen - auch in Ihrem eigenen Interesse - möchten wir Ihnen bereits an dieser Stelle danken.

Mit kollegialen Grüßen  
Für die Projektgruppe



Franz-Josef Klar

**Antwortcoupon**

Ich bin zur Mitarbeit bereit: ja ( ), nein ( )

Name:

Vorname:

Arzt/Ärztin ( )

Dipl.-Psychologe/in ( )

Anschrift (wenn möglich mit neuer Postleitzahl):

Telefon:

Abschluß der Ausbildung zum Psychoanalytiker (DGIP) im Jahre: .....

An welchem Institut haben Sie Ihre Ausbildung abgeschlossen? .....

Falls Sie zur Mitarbeit bereit sind: arbeiten Sie in eigener Praxis? ja ( )  
in einer Institution? ja ( ), wo anders? .....

Wieviele Patienten haben Sie zur Zeit in psychoanalytischer  
Behandlung? .....

Deutsche Gesellschaft für Psychoanalyse,  
Psychotherapie,  
Psychosomatik und Tiefenpsychologie e. V.

Vorsitzender:

Prof. Dr. med. Ulrich Streeck  
Tiefenbrunn - Krankenhaus für Psychotherapie  
und psychosomatische Medizin  
des Landes Niedersachsen  
3405 Rosdorf 1-Tiefenbrunn  
Tel. ~~0551 780 01 6~~ Fax ~~0551 780 27 70~~

Stellvertreter:

Prof. Dr. phil. Hans-Volker Werthmann  
Diplom-Psychologe  
Theodorenstraße 12 A, 6200 Wiesbaden  
Tel. ~~0611 787 12 10~~

Dr. med. Karin Bell  
Brucker Mauspfad 601  
5000 Köln 91  
Tel. ~~0221 43 05 20 10 42 07~~  
Fax über ~~0221 43 40 37~~

Fachpsychologen-Referat:

Dipl.-Psych. Ellen Bruckmayer  
Birkenstraße 1, 8133 Feldafing  
Tel. ~~089 37 732 47~~

Geschäftsführer / Justitiar:

Hochsachwalt Holger Schildt  
Johannisbollwerk 20, 2000 Hamburg 11  
Tel. ~~043 1 810 20 13~~ Fax ~~043 1 810 40 00~~

An die  
Vorsitzenden der Institute  
in der DGPT

An die  
Mitglieder des Vorstandes der DGPT

An den  
Vorsitzenden der DGAP  
Herrn Dr. Rudolf Müller

An den  
Vorsitzenden der DPG  
Herrn Prof. Dr. M. Ermann

An den  
Vorsitzenden der DPV  
Herrn Dr. C. Nedelmann

18. Sept. 1992

Liebe Kolleginnen, liebe Kollegen,

heute wende ich mich an Sie, um Sie auf eine Entwicklung aufmerksam zu machen, die nach meiner Einschätzung unserer besonderen Aufmerksamkeit und Wachsamkeit bedarf:

Wie Sie wahrscheinlich wissen, befaßt sich Herr Grawe in Bern seit einigen Jahren mit metaanalytischen Untersuchungen, die Aufschluß über die Frage der Effektivität verschiedener Psychotherapieverfahren geben sollen. Ergebnisse von Grawes Studien sind bisher in einer Reihe von Aufsätzen publiziert worden, und sie haben auch in das Gutachten zum Psychotherapeutengesetz Eingang gefunden; eine Veröffentlichung in Buchform soll demnächst erscheinen.

Eine etwas umfangreichere Arbeit von Grawe ist kürzlich auch in der Psychologischen Rundschau 1992, 43, S. 132-162 erschienen. Ich möchte diese Arbeit Ihrer besonderen Aufmerksamkeit empfehlen. Wie die Graweschen Arbeiten in der nicht-fachlichen Öffentlichkeit aufgenommen werden, davon hat ein Artikel in der "ZEIT", den Sie wahrscheinlich kennen, einen Vorgeschmack gegeben.

Die Studien von Grawe belegen angeblich nicht nur die völlige therapeutische Wirkungslosigkeit der Psychoanalyse bei psychosomatischen Krankheiten, sondern darüberhinaus, daß "alle mit der Langzeitanalyse verbundenen Hoffnungen gescheitert" sind.

Wir werden wohl damit zu rechnen haben, daß uns die angeblich gesicherten Ergebnisse der Graweschen Arbeiten schon bald von Kostenträgern, kassenärztlichen Vereinigungen u.a. mit dem Versuch vorgehalten werden, sie gegen uns zu verwenden. Auf diesem Hintergrund halte ich es für dringend erforderlich, daß wir uns detailliert und fachlich kompetent mit den Arbeiten von Grawe auseinandersetzen, um deren Einseitigkeit und Fragwürdigkeit, dort wo sie sich erweist, aufzeigen und kritisch dazu Stellung nehmen zu können. Entsprechende Stellungnahmen werden wir auch im berufspolitischen Feld benötigen.

Angesichts der hohen Relevanz für uns Psychoanalytiker und für unsere psychoanalytischen Ausbildungsinstitute bitte ich Sie sehr nachdrücklich darum, sich dieser Thematik anzunehmen, und ich bitte Sie darüberhinaus, die Ergebnisse Ihrer kritischen Auseinandersetzungen auch dem geschäftsführenden Vorstand für dessen Arbeit und für die zu erwartenden berufspolitischen Auseinandersetzungen zur Verfügung zu stellen.

Mit besten Grüßen bleibe ich  
Ihr

A handwritten signature in dark ink, appearing to read 'U. Streeck', written in a cursive style.

(Prof. Dr. U. Streeck)

Anlage



**4. Schreiben vom 30.06.1993 an alle mitwirkenden Psychoanalytiker/innen, incl. Untersuchungsdesign-Schema, Ablaufplan, Diagnoseschlüssel und Stellungnahme zu Bedenken**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. 0221/7408887

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleussner  
Dipl.-Pädagogin Monika Weber

vielen Dank für Ihre Antwort auf unser Schreiben vom April 93.

Wie angekündigt, möchten wir Sie heute über das weitere Procedere sowie über das Design und den Ablauf informieren. Darüber hinaus möchten wir auf einige kritische Rückmeldungen zu unserem ersten Schreiben eingehen.

Sie gehören entweder zur Gruppe der Kollegen/innen, die zur Mitarbeit bereit ist, oder Sie gehören zu der Gruppe, die zwar nicht abgeneigt ist, bei der Untersuchung mitzuwirken, jedoch Bedenken äußerte. Wir greifen diese Bedenken deswegen in der beigefügten Stellungnahme auf.

Falls Sie die Ihnen zugesandten Unterlagen nicht zurücksenden, gehen wir davon aus, daß Sie an der Untersuchung mitwirken. Nachdem der Patient die Fragebögen ausgefüllt an uns zurückgeschickt hat, warten wir ein halbes Jahr, bis wir Sie an den Re-Test erinnern. Sie werden dann von uns mit der Post wieder Testunterlagen erhalten. Wir übernehmen damit das Zeit-Management der Untersuchung und entlasten Sie von der Notwendigkeit, die Fristen der Re-Tests im Auge zu behalten.

Noch etwas ganz Wichtiges: Wir wären Ihnen sehr dankbar, wenn Sie mit so vielen Patienten in Untersuchungs- und Kontrollgruppe an der Untersuchung teilnehmen würden, wie Ihnen möglich erscheint. Sollten Sie also noch weitere Test-Materialien wünschen, z. B. weil Sie bereit sind, mit mehreren Patienten an der Untersuchung teilzunehmen, oder weil Sie bereit und in der Lage sind, der Untersuchung mehrere Wartepatienten zuzuführen, oder weil ein

.../2

Wartepatient aus unvorhergesehenen bzw. zwingenden Gründen doch vor Ablauf eines Jahres auf einen Therapieplatz kommt, dann geben Sie uns bitte sofort Nachricht; - wir reagieren dann kurzfristig und senden Ihnen neue Testunterlagen zu.

Als Anlage finden Sie

- Untersuchungsdesign
- Ablaufplan
- 4 Sätze Antwortbögen MMPI (mit einem Testheft)  
sowie 4 Gießener Beschwerdefragebögen
- 4 Freiumsschläge
- Diagnoseschlüssel (Verwendung siehe Ablaufplan)
- eine Stellungnahme zu den geäußerten Bedenken


Im Falle weiterer Rückfragen sollten wir diese telefonisch klären. Sie erreichen uns wie folgt:

Dieter Schleussner unter (02 11)4921976  
Franz Josef Klar, abends nach 21.00 Uhr  
unter (02 21)7408887

Wir hoffen auf ein gutes Gelingen.

Für die Forschungsgruppe

Köln, den 30.06.93



Franz Josef Klar

Anlage

# Forschungsgruppe Therapieeffektivität

## UNTERSUCHUNGSDESIGN

=====

### Stichprobe

UNTERSUCHUNGSGRUPPE: Patienten vor bzw. in analytischer Einzeltherapie

KONTROLLGRUPPE: Patienten für analytische Einzeltherapie vorgesehen,  
Beginn der Behandlung nach 1 Jahr Wartezeit

### Größe der Stichprobe

Untersuchungsgruppe      N = 200 Patienten  
Kontrollgruppe            N = 100 Patienten

### Untersuchungsinstrumente

- \* Minnesota Multiphasic Personality Inventory (MMPI), Kurzform
- \* Gießener Beschwerdebogen (GBB)
- \* Veränderungsfragebogen des Erlebens und Verhaltens (VEV)

	Untersuchungsgruppe		Kontrollgruppe
	Untersuchungszeitpunkt	Untersuchungsinstrument	Untersuchungszeitpunkt
1.	bei Beginn der analytischen Einzeltherapie, möglichst zwischen Erstgespräch und Ende der probatorischen Sitzungen	* MMPI, Kurzform * GBB	nach Erstgespräch
2.	nach 6 Monaten anal. Einzeltherapie	* MMPI, Kurzform * GBB * VEV	nach 6 Monaten Wartezeit
3.	nach 12 Monaten anal. Einzeltherapie	* MMPI, Kurzform * GBB * VEV	nach 12 Monaten Wartezeit
4.	Fortsetzung im 1-Jahrestakt vorgesehen bis 1 Jahr nach Behandlungsende	Kontrollgruppe wird zur 2. Untersuchungsgruppe Fortsetzung der Untersuchung im 1-Jahrestakt bis 1 Jahr nach Behandlungsende	

# Forschungsgruppe Therapieeffektivität

## A B L A U F P L A N =====

UNTERSUCHUNGSGRUPPE		KONTROLLGRUPPE
Auswahl von Pat. für analyt. ET	1.	Auswahl von Pat. für zukünftige analyt. ET
—	2.	Pat. Therapieplatz für spät. in 1 Jahr zusichern. Pat. sichert Teilnahme an Forschung zu.
Auf GBB tragen Sie bitte ein: Chiffre, Geschlecht, Anzahl bislang durchge- führter Sitzungen, Diagnose (DNR, Diagnoseschlüssel benutzen), Untersuchungsgruppe (UG) ankreuzen, Datum	3.	Auf GBB tragen Sie bitte ein: Chiffre, Geschlecht, Anzahl bislang durchge- führter Sitzungen, vorläufige Diagnose (DNR, Diagnoseschlüssel benutzen), Kontrollgruppe (KG) ankreuzen, Datum
Übergabe Testunterlagen (MMPI-Testheft mit Antwortbogen, GBB), Anschrift sowie blauen Umschlag an Pat.	4.	Übergabe Testunterlagen (MMPI-Testheft mit Antwortbogen, GBB), Anschriften sowie blauen Umschlag an Pat.
Beantwortung der Fragen durch Pat. möglichst in der Praxis, z.B. nach Therapiesitzung oder bei Pat. zu Hause	5.	Beantwortung der Fragen durch Pat. möglichst in der Praxis, z.B. nach Therapiesitzung oder bei Pat. zu Hause.
Pat. kuvertiert Antwortbögen ein und versendet sie in beigelegtem an uns adressierten blauen Umschlag.	6.	Pat. kuvertiert Antwortbögen ein und versendet sie in beigel. an uns adressierten blauen Umschlag.
Pat. gibt MMPI-Testheft an Sie zurück. Ablage des Testheftes bis Einsatz bei nächstem Pat.	7.	Pat. gibt MMPI-Testheft an Sie zurück. Ablage des Testheftes bis Einsatz bei nächstem Pat.
Nachfrage bei Pat., falls Rückgabe des Testheftes ausgeblieben: Sind Tests beantwortet und an Forschungsgruppe verschickt? Ggfs. Motivationshilfen geben.	8.	Nachfrage bei Pat., falls Rückgabe des Test- heftes ausgeblieben: Sind Tests beantwortet und an Forschungsgruppe verschickt? Ggfs. Motivationshilfen geben.
Unabhängig von der Anzahl der durchgeführten Sitzungen: Sie erhalten ca. 1/2 Jahr nach Behandlungs- beginn MMPI-Antwortbogen, GBB, VEV, Schreiben an Pat., blauen Umschlag	9.	Ca. 1/2 Jahr nach der ersten Befragung erhalten Sie von uns MMPI-Antwortbogen, GBB, VEV, Schreiben an Pat., blauen Umschlag.
Tragen Sie bitte auf GBB ein: Anzahl bislang durchgeführter Sitzungen und Diagnose (DNR, Diagnoseschlüssel benutzen)	10.	—
Übergabe Testunterlagen wie (4) und VEV	11.	Übergabe Testunterlagen wie (4) und VEV
Fragenbeantwortung wie (5)	12.	Fragenbeantwortung wie (5)
Versenden der Antwortbögen wie (6)	13.	Versenden der Antwortbögen wie (6)
Rückgabe wie (7)	14.	Rückgabe wie (7)
Nachfrage wie (8)	15.	Nachfrage wie (8)

**A) Diagnoseschlüssel nach vorherrschenden Strukturmerkmalen:**

Diagnose-Nr.  
(DNR)

- |    |                           |
|----|---------------------------|
| 1. | Hysterische Neurose       |
| 2. | Zwangsneurose             |
| 3. | Angstneurose              |
| 4. | Depressive Neurose        |
| 5. | Schizoide Neurose         |
| 6. | Depersonalisationssyndrom |
| 7. | Borderline-Störung        |
| 8. | Narzisstische Störung     |
| 9. | Sonstige Störungen        |

**B) Kürzel für Eintrag in GBB:**

Untersuchungsgruppe = UG

Kontrollgruppe = KG

# Forschungsgruppe Therapieeffektivität

## STELLUNGNAHME zu den geäußerten BEDENKEN

### ad KONTROLLGRUPPE:

Viele Kollegen teilten uns Bedenken hinsichtlich der Bildung der Kontrollgruppe und der damit verbundenen Notwendigkeit mit, Patienten ein Jahr auf eine Warteliste zu setzen. Wir können die Bedenken im einzelnen wohl verstehen, doch so leid es uns tut, wir kommen um die Bildung der Kontrollgruppe nicht umhin, wenn wir uns nicht von vorne herein dem methodischen 'Verriß' aussetzen wollen. Es kommt uns deshalb von Anfang an auf ein methodisch völlig korrektes Untersuchungs-Setting an, auch wenn dies Probleme mit sich bringt. Alles andere wäre nichts weiter als Verschwendung von Kosten und Mühen, und das sollten wir uns alle ersparen. Zu einer experimentellen Untersuchung gehört nun einmal die Kontrollgruppe, denn nur so können wir bei ansonsten vergleichbaren Gruppenbedingungen die Untersuchungsvariable "therapeutische Intervention" hinreichend isolieren und später Aussagen machen.

Wir gehen davon aus, daß uns Psychoanalytikern und besonders uns Analytikern in der DGIP in Zukunft ein rauher Wind entgegenbläst. Der empirische Wirksamkeitsnachweis wird, wie auch im organmedizinischen Bereich, zur Grundbedingung für die Zulassung der Psychotherapie als anerkanntes Heilverfahren. Unter Wirksamkeit wird verstanden, daß die Patienten infolge der Anwendung des Heilverfahrens ihre Krankheit überwinden oder daß ihre Krankheitssymptome deutlich nachlassen. Unser Untersuchungsdesign entspricht nach Selg den für univariate Experimente üblichen Vorgehen, in dem eine unabhängige Variable (hier: psychotherapeutische Behandlung) in ihrem Einfluß auf die abhängige Variable (hier: Befindlichkeit) untersucht wird, wobei die unabhängige Variable in einer Kontrollgruppe auf dem Nullniveau eingeführt werden muß (hier: es findet keine Behandlung statt).

Wir möchten Sie deshalb bitten, Ihre möglicherweise vorhandenen Bedenken zurückzustellen und uns bei der Bildung der Gruppe "wartende Therapiebewerber" zu helfen.

Um weiteren Nachfragen zuvorzukommen: Auch die Kollegen, bei denen wir die Bedenken vor allem hinsichtlich der 'Wartepatienten' nicht ausräumen konnten, oder die einfach aus anderen Gründen keine Warteliste oder Wartepatienten haben, können sich zumindest mit ihren demnächst beginnenden analytischen Therapien an der Untersuchung beteiligen. Wir sind dann allerdings sehr darauf angewiesen, daß andere Kollegen kompensierend mehrere Wartepatienten zur Untersuchung anmelden. Deshalb an dieser Stelle der Hinweis an die Kollegen, die eine umfängliche Warteliste haben, daß wir Ihnen auf Mitteilung gerne genügend Testmaterial zukommen lassen.

**ad ANONYMISIERUNG/DATENSCHUTZ:**

Andere Bedenken bezogen sich auf die u. U. aus Datenschutzgründen unzureichende Chiffrierung mit Hilfe der KV-Chiffrierungen. Wir können diese Bedenken nicht teilen. Da dieses Verfahren bundesweit praktiziert wird, gehen wir von der rechtlichen Unbedenklichkeit aus.

**ad UNTERSUCHUNGSBEGINN:**

Eine weitere Nachfrage betraf den Zeitpunkt des Beginns der Untersuchung: Uns ist klar, daß nicht alle Kollegen zum Stichtag x mit einem Patienten Y eine analytische Therapie beginnen werden. Es reicht, wenn Sie in der nächsten Zeit, in den nächsten Monaten eine Therapie beginnen und diese dann in die Untersuchung einbringen. Es ist nur wichtig, daß Sie im Berufsalltag unser gemeinsames Projekt nicht vergessen, und Ihre/n nächsten Patienten auch tatsächlich in die Untersuchung einbringen, denn wir können Sie natürlich nicht ständig erinnern und sind auf Ihre Initiative angewiesen.



## **5. Schreiben vom 30.06.1993 an alle nicht antwortenden Psychoanalytiker/innen**

## Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. 0221/7408887

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleusener  
Dipl.-Pädagogin Monika Weber

Liebe Kollegin,  
lieber Kollege

Wir haben Sie im April 93 angeschrieben, um Sie zur Mitarbeit bei der Wirksamkeitsuntersuchung der von uns IP-Analytikern durchgeführten analytischen Einzeltherapien zu bewegen.

Leider haben wir bislang von Ihnen keine Antwort erhalten.

Wir möchten Sie eindringlich bitten, die relativ geringe Mühe auf sich zu nehmen und sich an der Untersuchung zu beteiligen, die ja u.a. auch das Ziel hat, Ihre berufliche Existenz zu sichern.

Die Notwendigkeit der Untersuchung kann nicht ernstlich in Frage gestellt werden - es gilt, unsere Reputation und unsere zukünftige Beteiligung an der psychotherapeutischen Versorgung der Bevölkerung durch 'hieb- und stichfeste' empirische Forschung abzusichern.

Bitte geben Sie uns mit dem Ihnen bereits vorliegenden Freiumschlag und Antwortcoupon eine Rückmeldung.

Mit freundlichen Grüßen

Köln, den 30.06.93

*F. J. Klar*

## **6. Schreiben vom 21.04.1994 an alle mitwirkenden Psychoanalytiker/innen**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. 0221/7408887

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleussner  
Dipl.-Pädagogin Monika Weber

Liebe Kollegin,  
lieber Kollege

zunächst möchten wir uns nochmal für Ihre Mitarbeit an unserer Untersuchung zur Therapieeffektivität bedanken. Wir wenden uns heute jedoch mit der Bitte an Sie, weitere Patienten in die Untersuchung einzubringen.

Bislang besteht unsere Untersuchungsgruppe aus 90, die Kontrollgruppe aus 10 Patienten/innen. Wir sind zuversichtlich, daß wir die angestrebte Zielgröße von 200 Patienten in der Untersuchungsgruppe erreichen können, wenn Sie uns weiterhin tatkräftig unterstützen.

Was die Kontrollgruppe angeht, so sieht die Situation weniger günstig aus: 10 von den geplanten und auch erforderlichen Anzahl von 100 Patienten, das ist einfach noch zu wenig, um aussagekräftige Ergebnisse zu erhalten. Deshalb unsere Bitte an Sie: Sagen Sie zumindest einem oder zwei Patienten/innen, wenn Sie momentan keinen Platz für eine analytische Einzeltherapie frei haben, daß Sie ihm/ihr nach 1/2 oder 1 Jahr einen Platz zusichern. Wir sind uns im klaren darüber, daß diese Prozedur für Sie mit einigen Umständen verbunden ist. Keiner nimmt gerne Mühen auf sich, wenn sie sich vermeiden lassen. Bedenken Sie jedoch bitte, daß Sie einen Beitrag zu Ihrer eigenen Zukunftssicherung leisten. Denn mit der bevorstehenden Verabschiedung des Psychotherapeuten - Gesetzes sind wir keineswegs "aus dem Schneider". Der Gesetzgeber schreibt die Erbringung von anerkannten Wirksamkeitsnachweisen vor, bislang können wir dem nur mit unserer subjektiven Gewißheit über die Qualität unserer Therapie begeben.

Wir hoffen deshalb auf Ihre weitere Bereitschaft, an der Untersuchung mitzuwirken. Aus Kostengründen können wir jedoch nicht beliebig viel Testmaterial verschicken. Teilen Sie uns deshalb brieflich oder telefonisch mit, für wieviele Patienten wir Ihnen noch Testmaterial zuschicken dürfen.

Mit freundlichen Grüßen  
für die Forschungsgruppe



F.J. Klar

Köln, den 21.04.94

## **7. Schreiben vom September 1994 an alle mitwirkenden Psychoanalytiker/innen**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. 0221/7408887

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleussner  
Dipl.-Pädagogin Monika Weber

Sehr geehrte Frau Kollegin,  
sehr geehrter Herr Kollege,

nach einem Jahr intensiven Bemühens um unser Forschungsvorhaben haben wir nun berechnete Aussicht, das anvisierte Ziel von 200 Untersuchungspatienten und 100 Kontrollpatienten zu erreichen.

Um schneller ans Ziel zu kommen, erscheint es uns wenig sinnvoll, die Kollegen/innen nochmals anzusprechen, die wir bisher trotz mehrfacher Versuche nicht zur Mitarbeit motivieren konnten.

Stattdessen möchten wir uns noch einmal an die 53 Kollegen/innen wenden, die sich bisher mit ihren Patienten zur Mitarbeit bereitgefunden haben.

Wenn jeder der 53 Kollegen/innen nur noch 2 Patienten für die Untersuchungsgruppe und 1 Patienten für die Kontrollgruppe einbringt, kann die erste wichtige Phase der Datensammlung beendet werden. (Da viele Kollegen/innen keine 'Warteliste' haben, möchten wir darum bitten, daß andere Kollegen/innen vielleicht mehr als nur einen weiteren Wartepatienten einbringen.)

Diese Rechnung geht aber nur auf, wenn sich wirklich jeder der 53 Kollegen/innen noch einmal aufrufft und mitmacht. Jeder einzelne Ausfall wiegt schwer. Deshalb zählen wir auf jeden einzelnen, auch auf Sie.

Mit vielen kollegialen Grüßen  
für die Forschungsgruppe

Köln, den 21.09.94



In der Hoffnung auf Ihre Mitarbeit vorsorglich als Anlage: 3 komplette Testsets (das Testheft MMPI müßte Ihnen noch vorliegen.)

## **8. Schreiben vom Oktober 1994 an alle mitwirkenden Psychoanalytiker/innen**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. 0221/7408887

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleussner  
Dipl.-Pädagogin Monika Weber

Liebe Kolleginnen,  
liebe Kollegen,

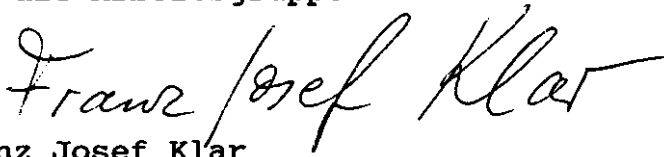
wir bedanken uns für die bislang eingegangenen  
Unterlagen.

## Folgende Probleme sind aufgetreten:

1. GBB (Gießener Beschwerdefragebogen) und MMPI erreichten uns oftmals mit getrennter Post. Deshalb konnten wir das Testmaterial nicht zuordnen, weil der MMPI-Fragebogen (anders als der GBB) von uns keine Datenzeile erhalten hat. Unsere dringende Bitte: Weisen Sie die Patienten noch einmal darauf hin, daß sie beide Unterlagen ausfüllen und zusammen zurückschicken sollen, und zwar in dem von uns zur Verfügung gestellten Couvert. Zur Sicherheit bitten wir Sie weiterhin, vor Aushändigung des Testmaterials an die Patienten auf dem MMPI-Fragebogen die Chiffre-Nummer des Patienten oben rechts hinzuschreiben, um so auf alle Fälle die Zuordnung sicherzustellen.
2. Anlässlich verschiedener Anfragen noch ein Hinweis für die Patienten: Bei der Bearbeitung des MMPI ist der Antwortbogen hinter die letzte Seite des Testheftes einzulegen.
3. Aufgrund verschiedener Rückfragen zur Bildung der Kontrollgruppe möchten wir Sie noch einmal darauf hinweisen und auch darum bitten, daß Sie eine unter Umständen bereits in Ihrer Praxis bestehende Warteliste in unsere Untersuchung als Kontrollgruppe einbringen.

Auf weitere gute Zusammenarbeit

für die Arbeitsgruppe



Franz Josef Klar



**9. Anschreiben an alle Patienten der Untersuchungsgruppe und der Kontrollgruppe zum Messzeitpunkt t1**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleussner  
Dipl.-Pädagogin Monika Weber

Sehr geehrte Dame,  
sehr geehrter Herr,

Wir bitten Sie um Unterstützung bei einem wichtigen Forschungsvorhaben zur Untersuchung der Wirksamkeit psychotherapeutischer Behandlung. Dazu gehört, daß wir Sie heute und zu späteren Zeitpunkten (nach 1/2 Jahr, nach 1 Jahr, nach 2 Jahren) befragen. Wir sind uns bewußt, daß für Sie damit ein erheblicher Zeitaufwand verbunden ist. Für die erfolgreiche Durchführung unseres Projektes sind wir aber auf Ihre Hilfe angewiesen.

Wenden Sie sich nun bitte den beiliegenden Fragebogen zu:  
Achten Sie bitte genau auf die Anweisungen für die Beantwortung der Fragen.  
Die vollständig angekreuzten Antwortbögen senden Sie bitte im beigelegten Umschlag an uns zurück.  
Wir sichern Ihnen vollkommene Anonymität zu.  
Das Testheft MMPI schicken bzw. geben Sie bitte Ihrem Therapeuten/Ihrer Therapeutin zurück.

Am besten, Sie nehmen sich jetzt gleich Testheft und Fragebögen vor und beantworten die Fragen zügig und vor allem ungestört und unbeeinflußt durch andere.  
Nicht vergessen: Datum der Beantwortung auf dem Bogen GBB eintragen!

Vielen Dank für Ihre freundliche Unterstützung.

Für die Forschungsgruppe



**10. Anschreiben an alle Patienten der Untersuchungsgruppe und der Kontrollgruppe zum Messzeitpunkt t2, einem halben Jahr nach t1**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. 0221/7408887

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schieussner  
Dipl.-Pädagogin Monika Weber

Sehr geehrte Dame,  
sehr geehrter Herr,

Sie befinden sich z.Z. in einer Psychotherapie, bzw. Sie warten auf einen zugesagten Therapieplatz, und können sich wahrscheinlich erinnern, daß Sie vor ca. einem halben Jahr zwei Fragebögen ausgefüllt haben, die wir Ihnen über Ihren Therapeuten zugeleitet haben.

Wir möchten die Gelegenheit nutzen, uns an dieser Stelle zunächst für Ihre freundliche Mitarbeit bei einem wichtigen Forschungsvorhaben zur Wirksamkeit psychotherapeutischer Behandlung bedanken.

Wie wir Ihnen in unserem ersten Anschreiben bereits mitgeteilt haben, gehört es zu der Untersuchung, daß wir Sie in Abständen von einem halben Jahr immer wieder erneut befragen müssen. Nur so können wir die Wirksamkeit der Therapie feststellen.

Wir sind uns bewußt, daß damit für Sie ein erheblicher Zeitaufwand verbunden ist. Für die erfolgreiche Durchführung des Projekts ist dies jedoch der einzig gangbare Weg. Wir sind hierbei auf Ihre Hilfe und Bereitschaft zur Mitarbeit dringend angewiesen.

Wir legen Ihnen deshalb heute erneut 3 Fragebögen vor. Zwei der Fragebögen kennen Sie bereits, ein dritter Fragebogen ist hinzugekommen.

Wenden Sie sich nun bitte den beiliegenden Fragebögen zu: Achten Sie bitte genau auf die Anweisungen für die Beantwortung der Fragen. Die vollständig ausgefüllten Antwortbögen senden Sie bitte alle zusammen in dem beigefügten Umschlag an uns zurück. Sie dürfen die Bögen hierzu knicken. Bitte schicken Sie uns die Bögen nicht einzeln zu, da wir sie sonst nicht als zusammengehörig zuordnen können. Wir sichern Ihnen wie beim ersten mal wieder vollkommene Anonymität zu. Das Testheft MMPI schicken bzw. geben Sie bitte Ihrem Therapeuten/Ihrer Therapeutin zurück.

Am besten, Sie füllen die Bögen jetzt gleich aus und beantworten die Fragen zügig und vor allem ungestört und unbeeinflußt durch andere.

Nicht vergessen: Auf dem Bogen GBB das Datum der Beantwortung eintragen.

Vielen Dank für Ihre freundliche Unterstützung.

Für die Forschungsgruppe



**Haben Sie zwischenzeitlich an anderen psychotherapeutischen Verfahren teilgenommen ?**

**Falls ja, wieviele Stunden ?**

	<b>Ja</b>	<b>nein</b>	<b>Std.</b>
<b>Einzeltherapie</b>	<b>O</b>	<b>O</b>	_____
<b>Gruppentherapie</b>	<b>O</b>	<b>O</b>	_____

**Haben Sie zwischenzeitlich an anderen psychotherapeutischen Verfahren teilgenommen ?**

**Falls ja, wieviele Stunden ?**

	<b>Ja</b>	<b>nein</b>	<b>Std.</b>
<b>Einzeltherapie</b>	<b>O</b>	<b>O</b>	_____
<b>Gruppentherapie</b>	<b>O</b>	<b>O</b>	_____

**11. Anschreiben an alle Patienten der Untersuchungsgruppe und der Kontrollgruppe zum Messzeitpunkt t3, einem halben Jahr nach t2 und einem Jahr nach t1**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. 0221/7408887

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleussner  
Dipl.-Pädagogin Monika Weber

Sehr geehrte Dame,  
sehr geehrter Herr,

es ist wieder ein halbes Jahr vergangen, seitdem wir Sie das letzte Mal gebeten haben, unsere Fragebögen zu beantworten.

Wie Sie ja bereits wissen, versuchen wir Ihre Befindlichkeit in bestimmten Intervallen zu erfassen, so daß wir Sie heute bitten möchten, die Ihnen bereits bekannten Fragebögen erneut auszufüllen.

Achten Sie bitte genau auf die Anweisungen für die Beantwortung der Fragen. Die vollständig beantworteten Antwortbögen senden Sie bitte alle zusammen in dem beigelegten Umschlag an uns zurück. Sie dürfen die Bögen hierzu knicken. Bitte schicken Sie uns die Bögen nicht einzeln zu, da wir sie sonst nicht als zusammengehörig zuordnen können.

Wir sichern Ihnen wieder vollkommene Anonymität zu.

Das Testheft MMPI schicken bzw. geben Sie bitte Ihrem Therapeuten/Ihrer Therapeutin zurück.

Am besten, Sie füllen die Bögen jetzt gleich aus und beantworten die Fragen zügig und vor allem ungestört und unbeeinflußt durch andere.

Nicht vergessen: Auf dem Bogen GBB das Datum der Beantwortung eintragen.

Wir möchten nicht versäumen, uns nochmals recht herzlich für Ihre Mühe zu bedanken.

Mit freundlichen Grüßen

Für die Forschungsgruppe



**12. Anschreiben an alle Patienten der Untersuchungsgruppe zum Messzeitpunkt t4 und t5**



# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. 0221/7408887

Leitung:  
Prof. Dr. Günter Heisterkamp  
Wissenschaftliche Beratung:  
Dr. Rolf Dellen

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleussner  
Dipl.-Pädagogin Monika Weber

Sehr geehrte Dame,  
sehr geehrter Herr,

es ist wieder ein halbes Jahr vergangen, seitdem wir Sie das letzte Mal gebeten haben, unsere Fragebögen zu beantworten.

Wie Sie ja bereits wissen, versuchen wir Ihre Befindlichkeit in bestimmten Intervallen zu erfassen, so daß wir Sie heute bitten möchten, die Ihnen bereits bekannten Fragebögen erneut auszufüllen.

Achten Sie bitte genau auf die Anweisungen für die Beantwortung der Fragen. Die vollständig beantworteten Antwortbögen senden Sie bitte alle zusammen in dem beigefügten Umschlag an uns zurück. Sie dürfen die Bögen hierzu knicken. Bitte schicken Sie uns die Bögen nicht einzeln zu, da wir sie sonst nicht als zusammengehörig zuordnen können.

Wir sichern Ihnen wieder vollkommene Anonymität zu.

Das Testheft MMPI schicken bzw. geben Sie bitte Ihrem Therapeuten/Ihrer Therapeutin zurück.

Am besten, Sie füllen die Bögen jetzt gleich aus und beantworten die Fragen zügig und vor allem ungestört und unbeeinflußt durch andere.

Nicht vergessen: Auf dem Bogen GBB das Datum der Beantwortung eintragen.

Wir möchten nicht versäumen, uns nochmals recht herzlich für Ihre Mühe zu bedanken.

Mit freundlichen Grüßen

Für die Forschungsgruppe



**13. Schreiben vom 24.08.1999 an alle mitwirkenden Psychoanalytiker/innen zur Vorbereitung der Katamnese**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
~~Tel. 0221/7408887~~

Leitung:  
Prof. Dr. Günter Heisterkamp  
~~Wissenschaftliche Beratung:~~  
~~Dr. Rolf Dellen~~

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schlaussner  
Dipl.-Pädagogin Monika Weber

Sehr geehrte Kollegin,  
sehr geehrter Kollege,

Köln, den 24.08.99

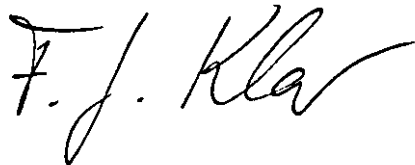
Sie haben über lange Zeit unsere Forschungsarbeit unterstützt, indem Sie ein oder mehrere Patienten in die Untersuchung eingebracht und Fragebögen von uns an Sie weitergegeben haben. Dafür dürfen wir uns an dieser Stelle recht herzlich bei Ihnen bedanken.

Wir möchten nun noch eine abschließende Katamneseuntersuchung durchführen. Dazu ist es notwendig zu wissen, ob die von Ihnen in die Untersuchung eingebrachten Patienten die Behandlung beendet haben und wenn ja, zu welchem Zeitpunkt. Bitte senden Sie zu diesem Zweck den beiliegenden Bogen ausgefüllt an uns zurück oder faxen Sie unter Nr.: 0221-765778.

Für die Patienten, die die Therapie beendet haben, schicken wir Ihnen dann in Kürze noch ein letztes Mal Fragebögen zur Weiterleitung zu.

Vielen Dank für Ihre Mitarbeit.

Für die Forschungsgruppe



**14. Katamneseanschreiben an alle Psychoanalytiker/innen, mit Anschreiben an die Patienten und mit Fragebögen zur Katamnese**

# Forschungsgruppe Therapieeffektivität

Sekretariat:

Kirburgerweg 21

50767 Köln

Tel.: ~~02696 930099~~

Leitung:

Prof. Dr. Günter Heisterkamp

Wissenschaftliche Mitarbeiter:

Dipl.-Psychologe Franz-Josef Klar

Dipl.-Psychologe Dieter Schleussner

Dipl.-Pädagogin Monika Weber

Sehr geehrte Frau Kollegin,  
sehr geehrter Herr Kollege,

Sie haben das Forschungsprojekt Therapieeffektivität unterstützt, indem Sie einen/einige Patienten in die Untersuchung eingebracht haben.

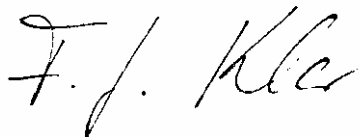
Wir möchten sie heute bitten, diesen Patienten einige Fragebögen zukommen zu lassen, mit denen wir das Befinden der Patienten nach Beendigung der Behandlung erheben wollen. Wir wären Ihnen sehr dankbar, wenn Sie uns bei dieser doch sehr wichtigen katamnestischen Befragung noch einmal unterstützen würden.

Die Patienten schicken dann die Fragebögen in dem vorbereiteten Umschlag direkt an uns zurück.

Versehen Sie bitte den von uns frankierten Umschlag mit der Patientenadresse. In diesen stecken Sie bitte das von Ihnen unterschriebene Begleitschreiben und den Umschlag mit den Patientenunterlagen und senden diesen Umschlag an die/den Patientin/en.

Vielen Dank für Ihre Mitarbeit.

Die Forschungsgruppe



Sehr geehrte/r Frau/Herr .....,

die Forschungsgruppe Therapieeffektivität ist noch einmal an mich herangetreten mit der Bitte, Ihnen Fragebögen zu Ihrem Befinden nach Beendigung Ihrer Therapie zukommen zu lassen. Ich unterstütze das Anliegen der Forschungsgruppe und möchte Sie deshalb bitten, die Fragebögen auszufüllen. Sie benötigen dazu lediglich einige Minuten. Die Anonymität der Befragung ist gesichert.

In der Hoffnung, daß es Ihnen gut geht, bedanke ich mich bei Ihnen für Ihre Unterstützung und verbleibe

mit freundlichen Grüßen

# Forschungsgruppe Therapieeffektivität

Sekretariat:

Kirburgerweg 21

50767 Köln

Tel.: ~~02696 990099~~

Leitung:

Prof. Dr. Günter Heisterkamp

Wissenschaftliche Mitarbeiter:

Dipl.-Psychologe Franz-Josef Klar

Dipl.-Psychologe Dieter Schleussner

Dipl.-Pädagogin Monika Weber

Sehr geehrte Dame,  
sehr geehrter Herr,

Sie haben in der Zeit Ihrer psychotherapeutischen Behandlung unsere Forschungsgruppe unterstützt, indem Sie zu mehreren Zeitpunkten während Ihrer Behandlung Fragebogen ausgefüllt haben. Damit haben Sie ein wichtiges Forschungsvorhaben gefördert, mit dem die Wirksamkeit psychotherapeutischer Behandlung wissenschaftlich untersucht wird.

Heute kommen wir noch einmal mit einer Bitte auf Sie zu: Wir möchten gerne wissen, wie es Ihnen heute, einige Zeit nach Beendigung Ihrer Psychotherapie geht und wie Sie heute die Zeit Ihrer Behandlung einschätzen.

Nehmen Sie bitte dazu die beiliegenden Fragebögen und beantworten Sie der Reihe nach. Selbstverständlich bleibt auch diese Befragung anonym. Stecken Sie bitte die ausgefüllten Bögen in den vorbereiteten Briefumschlag und senden diesen an uns zurück.

Vielen Dank für Ihre Mitarbeit.

Die Forschungsgruppe

**Chiffre:** \_\_\_\_\_

**letzte Therapiesitzung:**      **Monat** \_\_\_\_\_ **Jahr** \_\_\_\_\_

**Geschlecht:**

männlich                      ( )

weiblich                      ( )

**Familienstand:**

1) ledig                      ( )

2) verheiratet                      ( )

3) getrennt lebend                      ( )

4) geschieden                      ( )

5) verwitwet                      ( )

6) wieder verheiratet                      ( )

**Anzahl der eigenen**

**Kinder:**                      ( )

(0) keine                      ( )

**Alter in Jahren:**

15 - 20                      ( )

21 - 30                      ( )

31 - 40                      ( )

41 - 50                      ( )

51 - 60                      ( )

61 und älter                      ( )

**Schulabschluß:**

1) Hauptschule                      ( )

2) Mittlere Reife                      ( )

3) Abitur                      ( )

4) kein Abschluß                      ( )

**Gegenwärtiger Beruf oder zuletzt ausgeübter Beruf:**

01)    Freier Beruf (Arzt, Anwalt usw.)                      ( )

        Leiter von Unternehmen                      ( )

02)    Höherer Beamter, leitender Angestellter                      ( )

03)    Beamter, Angestellter                      ( )

04)    Selbständiger Gewerbetreibender                      ( )

        Landwirt                      ( )

05)    Facharbeiter                      ( )

06)    Arbeiter                      ( )

07)    Hausfrau, Hausmann, mithelfender Familienangehöriger                      ( )

08    Schüler, Auszubildender, Student                      ( )

09)    Rentner, Pensionär                      ( )

10)    ohne Beruf                      ( )

11)    arbeitslos                      ( )



Dieser Fragebogen erkundigt sich nach Ihrem Verhalten und Erleben gegenüber sich selbst. Bitte antworten Sie so, wie Sie wirklich denken oder fühlen. Die allererste Reaktion auf die jeweilige Frage ist meist die beste Antwort. Es gibt keine „guten“ oder „schlechten“ Antworten. Nur Ihre Meinung zählt. Es geht nicht um „wahr“ oder „falsch“ oder darum, was andere von Ihnen erwarten.

Kreuzen Sie nun bitte an, wie gut die jeweilige Frage beschreibt, wie Sie mit sich selbst umgehen.

**So gehe ich in meinen besten Zeiten mit mir um:**

	stimmt über- haupt nicht	0	1	2	3	4	5	6	7	8	9	10	stimmt voll und ganz
Unbekümmert und wie selbstverständlich tue und lasse ich, wonach mir der Sinn steht.		0	1	2	3	4	5	6	7	8	9	10	
Ohne einen Gedanken an die Folgen bin ich voller Ablehnung und Zerstörungswut mir gegenüber.		0	1	2	3	4	5	6	7	8	9	10	
Ich mag mich und gehe liebevoll mit mir um.		0	1	2	3	4	5	6	7	8	9	10	
Im Umgang mit mir selbst setze ich einiges daran, fürsorglich, achtsam und interessiert an meiner eigenen Entwicklung zu sein.		0	1	2	3	4	5	6	7	8	9	10	
Ich bestrafe mich durch Selbstvorwürfe, Selbstzweifel und Selbsterniedrigung.		0	1	2	3	4	5	6	7	8	9	10	
Ich nehme mich so wie ich bin, mit all meinen Stärken und Schwächen.		0	1	2	3	4	5	6	7	8	9	10	
Gedankenlos und rücksichtslos vernachlässige ich mich, manchmal so, als wäre ich ganz egal.		0	1	2	3	4	5	6	7	8	9	10	
Ich achte darauf, daß ich alles richtig mache und halte mich selbst unter genauer Beobachtung und Kontrolle.		0	1	2	3	4	5	6	7	8	9	10	

Nun stellen Sie sich bitte um und beurteilen Sie:

**So gehe ich in meinen schlechtesten Zeiten mit mir um :**

	stimmt über- haupt nicht	0	1	2	3	4	5	6	7	8	9	10	stimmt voll und ganz
Unbekümmert und wie selbstverständlich tue und lasse ich, wonach mir der Sinn steht.		0	1	2	3	4	5	6	7	8	9	10	
Ohne einen Gedanken an die Folgen bin ich voller Ablehnung und Zerstörungswut mir gegenüber.		0	1	2	3	4	5	6	7	8	9	10	
Ich mag mich und gehe liebevoll mit mir um.		0	1	2	3	4	5	6	7	8	9	10	
Im Umgang mit mir selbst setze ich einiges daran, fürsorglich, achtsam und interessiert an meiner eigenen Entwicklung zu sein.		0	1	2	3	4	5	6	7	8	9	10	
Ich bestrafe mich durch Selbstvorwürfe, Selbstzweifel und Selbsterniedrigung.		0	1	2	3	4	5	6	7	8	9	10	
Ich nehme mich so wie ich bin, mit all meinen Stärken und Schwächen.		0	1	2	3	4	5	6	7	8	9	10	
Gedankenlos und rücksichtslos vernachlässige ich mich, manchmal so, als wäre ich ganz egal.		0	1	2	3	4	5	6	7	8	9	10	
Ich achte darauf, daß ich alles richtig mache und halte mich selbst unter genauer Beobachtung und Kontrolle.		0	1	2	3	4	5	6	7	8	9	10	

Kreuzen Sie bitte die für Sie zutreffende Zahl an.

	ich halte dies für sehr unzutreffend				ich halte dies für sehr zutreffend	
1. Ich glaube, daß mein Therapeut mir geholfen hat.	1	2	3	4	5	6
2. Ich glaube, daß die Behandlung mir geholfen hat.	1	2	3	4	5	6
3. Ich habe einige neue Einsichten gewonnen.	1	2	3	4	5	6
4. Seit einiger Zeit fühle ich mich besser.	1	2	3	4	5	6
5. Ich sehe, daß ich die Probleme bewältigen kann, wegen derer ich zur Behandlung kam.	1	2	3	4	5	6
6. Ich hatte das Gefühl, mich auf den Therapeuten verlassen zu können.	1	2	3	4	5	6
7. Ich hatte das Gefühl, daß mich der Therapeut versteht.	1	2	3	4	5	6
8. Ich hatte das Gefühl, daß mein Therapeut möchte, daß ich meine Ziele erreiche.	1	2	3	4	5	6
9. Ich hatte das Gefühl, daß ich wie auch der Therapeut ernsthaft an einem Stang ziehen.	1	2	3	4	5	6
10. Ich glaube, daß ich und der Therapeut meine Probleme ähnlich gesehen und beurteilt haben.	1	2	3	4	5	6
11. Ich habe das Gefühl, daß ich mich jetzt selbst verstehen und mich mit mir auseinandersetzen kann.	1	2	3	4	5	6
	sehr gut				sehr schlecht	
12. Insgesamt schätze ich den Therapieerfolg ein als:	1	2	3	4	5	6

Beantworten Sie nun bitte noch die beiden folgenden Fragen.

1. Schätzen Sie bitte ein, wie häufig Sie sich in Ihrer Therapie gefreut haben:

*(Kreuzen Sie bitte die zutreffende Antwort an!)*

☐ ständig      ☐ sehr oft      ☐ hin u. wieder      ☐ selten      ☐ sehr selten      ☐ nie

2. Wie häufig haben Sie sich in der Zeit Ihrer Therapie gemeinsam mit Ihrer Therapeutin/Ihrem Therapeuten gefreut?

*(Kreuzen Sie bitte die zutreffende Antwort an!)*

☐ ständig      ☐ sehr oft      ☐ hin u. wieder      ☐ selten      ☐ sehr selten      ☐ nie

3. Schildern Sie nun bitte möglichst konkret eine freudige Situation aus Ihrer Therapie:

---

---

---

---

---

---

## **15. Anschreiben zur Nachfassaktion im Sommer 2000**

# Forschungsgruppe Therapieeffektivität

Sekretariat:  
Jesuitengasse 71  
50735 Köln  
Tel. ~~0221/7408897~~

Leitung:  
Prof. Dr. Günter Heisterkamp

Wissenschaftliche Mitarbeiter:  
Dipl.-Psychologe Franz-Josef Klar  
Dipl.-Psychologe Dieter Schleussner  
Dipl.-Pädagogin Monika Weber

Sehr geehrte/r Frau/Herr

wir möchten Sie bitten, uns für den/die von Ihnen  
in unsere Untersuchung eingebrachte/n Patientin/-ten  
noch die Auskunft zu geben, welchen Schulabschluß  
(Hauptschule/Mittlere Reife/Abitur/Studium) die/der  
Patient hat, sowie welchen Beruf sie/er bei Therapie-  
beginn ausübte.

Der Einfachheit halber sprechen Sie uns bitte auf  
Band unter ~~02696 930033~~ oder senden Sie ein Fax an  
~~02696 930033~~.

Die Chiffre des Patienten, zu dem uns die o.g. Infor-  
mation noch fehlt ist:

Vielen Dank für Ihre Bemühung.

Mit freundlichen Grüßen

